ISSN 0019-5588

Indian Journal of Pure & Applied Mathematics

DEVOTED PRIMARILY TO ORIGINAL RESEARCH IN PURE AND APPLIED MATHEMATICS

VOLUME 20/1
JANUARY 1989

INDIAN JOURNAL OF PURE AND APPLIED MATHEMATICS

Published monthly by the

INDIAN NATIONAL SCIENCE ACADEMY

Subscriptions

For India, Pakistan, Sri Lanka, Nepal, Bangladesh and Burma, Contact :

Associate Editor, Indian National Science Academy, Bahadur Shah Zafar Marg, New Delhi 110002, Telephone : 3311865, Telex : 31-61835 INSA IN.

For other countries, Contact :

M/s J. C. Baltzer AG, Scientific Publishing Company, Wettsteinplatz 10, CH-4058 Basel, Switzerland, Telephone : 61-268925, Telex : 63475.

The Journal is indexed in the Science Citation Index; Current Contents (Physical, Chemical & Earth Sciences); Mathematical Reviews; INSPEC Science Abstracts (Part A); as well as all the major abstracting services of the World.

Indian J. pure appl. Math., 20 (1) : 1-9, January 1989

ON THE AVERAGE NUMBER OF CROSSINGS OF AN ALGEBRAIC POLYNOMIAL

K. FARAHMAND

Department of Mathematical Statistics, University of Natal, King George V Avenue Durban 4001, *South Africa*

(*Received* 31 *August* 1987; *after revision* 30 *March* 1988)

The present paper provides an estimate of the expected number of crossings of a polynomial of degree n with the line $y = mx$ where the coefficients are independent normally distributed and m is a constant indepenpent of x. There are many known asymptotic estimates for the case of $m=0$. It is shown that the result is still valid even for $m \to \infty$ as long as $m = o(\sqrt{n})$.

1. INTRODUCTION

Let

$$P(x) = \sum_{i=0}^{n-1} a_i x^i \qquad \ldots(1.1)$$

where $a_0, a_1, a_2, \ldots, a_{n-1}$ is a sequence of random variables, and $N_m(a, b) \equiv N(a, b)$ be the number of real roots of the algebraic equation $P(x) = mx$ in the interval (a, b), where $m_n \equiv m$ is a constant independent of x, and multiple roots are counted only once. Kac[5] found that in the case of $m = 0$, and when the coefficients of (1.1) are independent normally distributed with mean zero and variance one, the mathematical expectation of the number of real roots, $EN(-\infty, \infty)$, is asymptotic to $(2/\pi)\log n$. From later works[3,6,8], there is ground to believe that whatever class of distributions for the coefficients we choose it will not greatly affect the result as long as $E(a_i) = 0$ $(i = 0, 1, 2, \ldots, n-1)$. Further a reduction to the expected number of real roots appears to occur in the work of Ibragimov and Maslova[4] and Sambandham[7] when they consider the cases of the coefficients having non zero means or being dependent. An asymptotic formula for the expected number of real roots for $P(x)=m$ is obtained by the author[2], which is, indeed, the number of times that $P(x)$ crosses a line parrallel to the x-axis. Here we study the number of times that $P(x)$ crosses a line which is not necessarily parallel to the x-axis. We state our result in the following theorem.

Theorem—If the coefficients of (1.1) are independent, standard normal random variables, then for any constant m such that (m^2/n) tends to zero the mathematical expectation of the number of real roots of the equation $P(x) = mx$ satisfies

$$EN\,(-1,1) \sim (1/\pi)\log(n/m^2) \quad \text{if } m \to \infty \text{ as } n \to \infty$$

$$EN\,(-1,1) \sim (1/\pi)\log n \qquad \text{if } m \text{ is bounded}$$

$$EN\,(-\infty, -1) = EN\,(1, \infty) \sim (1/2\pi)\ \text{long}\ n.$$

From the theorem it is interesting to note that for sufficiently large n we still obtain a sizeable number of crossings even when the line tends to be perpendicular to the x-axis (i. e. $m \to \infty$).

2. A Formula for the Number of Crossings

Let $\phi(x) = (2\pi)^{-1/2} \int_{-\infty}^{x} \exp(-y^2/2)\,dy$ and $\phi(x) = \phi'(x) = (2\pi)^{-1/2} \exp(-x^2/2)$, then by using the expected number of level crossings [(Cramer and Leadbetter[1], p. 285)] for our equation $P(x) - mx = 0$ we have

$$EN\,(a, b) = \int_a^b \beta^{1/2} \alpha^{-1/2} (1 - \mu^2)^{1/2} \phi\,(\omega/\alpha)^{1/2} [2\phi(\eta) + \eta\,\{2\phi(\eta) - 1\}]\,dx \ldots (2.1)$$

where

$$\alpha(x) \equiv \alpha = \text{Var}\,\{P(x) - mx\}, \quad \beta(x) \equiv \beta = \text{Var}\,[\{P(x) - mx\}']$$

$$\mu = \alpha^{-1/2} \beta^{-1/2} \text{Cov}\,[\{P(x) - mx\}, \{P(x) - mx\}'], \ \omega = E\,\{P(x) - mx\}$$

$$\eta = \beta^{-1/2} (1 - \mu^2)^{-1/2} \{\nu - \beta^{1/2} \mu\,\omega\,\alpha^{-1/2}\} \text{ and } \nu = E\,[\{P(x) - mx\}'].$$

Since the coefficients of $P(x)$ are independent random variables with mean zero and variance one we can easily find that

$$\omega = -mx, \qquad \alpha = \text{Var}\,\{P(x)\} = \sum_{i=1}^{n-1} x^{2i}$$

$$\beta = \text{Var}\,\{P'(x)\} = \sum_{i=1}^{n-1} i^2 x^{2i-2}, \ \nu = -m$$

$$\mu = \alpha^{-1/2} \beta^{-1/2} E\,\{P(x).\,P'(x)\} = \alpha^{-1/2} \beta^{-1/2} \sum_{i=1}^{n-1} i\,x^{2i-1} = \alpha^{-1/2} \beta^{-1/2} \gamma$$

(say)

and

$$\eta = m\,\alpha^{-1/2} (\alpha\beta - \gamma^2)^{-1/2} (\gamma x - \alpha).$$

Hence from (2.1) and since $\phi(x) = \frac{1}{2} + (2\pi)^{-1/2} \text{erf}(x)$ we have the extension of Kac-Rice formula[9].

$$EN\,(a, b) = \int_a^b (\Delta^{1/2}/\pi\alpha) \exp\{(-\alpha m^2 + 2m^2 \gamma x - \beta m^2 x^2)/2\Delta\}$$

(equation continued on p. 3)

$$+ \{ \mid m(\gamma x - \alpha) \mid / \pi\alpha^{3/2}\} \exp(-m^2 x^2/2\alpha) \operatorname{erf}$$

$$\{ \mid m(\gamma x - \alpha) \mid / \alpha^{1/2} \Delta^{1/2}\} \, dx$$

$$= \int_a^b I(x)\, dx = \int_a^b I_1(x)\, dx + \int_a^b I_2(x)\, dx \text{ (say)} \qquad \ldots(2.2)$$

where

$$\Delta(x) \equiv \Delta = \alpha\beta - \gamma^2.$$

3. Proof of the Theorem

First we consider the case when $m \to \infty$ as $n \to \infty$. For finding an upper limit of $EN(0, 1)$, let $0 \leqq a \leqq 1$ and divide $(0, 1)$ into two sub-intervals $(0, 1 - 1/n^a)$ and $(1 - 1/n^a, 1)$. For $0 \leqq x \leqq 1 - 1/n^a$ and all sufficiently large n we have

$$\gamma = \{(n-1)x^{2n+1} - nx^{2n-1} + x\}(1 - x^2)^{-2}$$

$$= x(1 - x^{2n})(1 - x^2)^{-2} + O\{n^{1+a}\exp(-2n^{1-a})\} \qquad \ldots(3.1)$$

and

$$\beta = (1 + x^2)(1 - x^{2n})(1 - x^2)^{-3} + O\{n^{2+a}\exp(-2n^{1-a})\}. \qquad \ldots(3.2)$$

From (3.1) and (3.2) we have

$$\Delta = (1 - x^{2n})(1 - x^2)^{-4} + O\{n^{2+2a}\exp(-2n^{1-a})\}. \qquad \ldots(3.3)$$

Now we choose $a = 1 - \log\log(n^{10})/\log n$. Then since

$$\exp(-2n^{1-a})\exp\{-2\log(n)^{10}\} = n^{-20}$$

for n sufficiently large, all the terms inside the $0\{\ \}$ will tend to zero and also $a \to 1$ which is necessary for obtaining our further result. Hence from (2.2), (3.1), and (3.3) for all sufficiently large n we have

$$\int_0^{1-1/n^a} I_1(x)\, dx = \int_0^{1-1/n^a} (\Delta^{1/2}\pi\alpha) \exp[\{-m^2(1 - x^2)/(1 - x^{2n})\}\{(1 - x^2)^2/2$$

$$- x^2(1 - x^2) + x^2(1 + x^2)/2\}\{1 + O(n^{2+a}$$

$$\times \exp(-2n^{1-a}))\}]\, dx. \qquad \ldots(3.4)$$

Now we note that for all x

$$(1 - x^2)^2/2 - x^2(1 - x^2) + x^2(1 + x^2)/2 = (1 - 3x^2 + 4x^4)/2$$

$$\geqq 7/32 > 1/5$$

and since from Kac[5] we have $\Delta^{1/2}/\alpha < (1 - x^2)^{-1}$, (3.4) gives

$$\int_0^{1-1/n^a} I_1(x)\, dx < (1/\pi) \int_0^{1-1/n^a} (1 - x^2)^{-1} \exp\{-m^2(1 - x^2)/5(1 - x^{2n})\}\, dx$$

(equation continued on p. 4)

$$< (1/\pi) \int_0^{1-1/n^a} (1 - x^2)^{-1} \exp\{-m^2 (1 - x^2)/5\}\, dx. \qquad \ldots(3.5)$$

Let $\lambda = m^2/5$, then from (3.5) and since $\exp\{-\lambda (1 - x)^2)\} < \{1 - \lambda (1 - x^2)^{-1}\}$ we have

$$\int_0^{1-1/n^a} I_1(x)\, dx < (1/\pi) \int_0^{1-1/n^a} (1 - x^2)^{-1} \{1 + \lambda (1 - x^2)\}^{-1} dx$$

$$= (1/\pi) \int_0^{1-1/n^a} [(1 - x^2)^{-1} - \lambda/\{1 + \lambda (1 - x^2)\}]\, dx$$

$$= (1/2\pi) \log \frac{2 - 1/n^a}{1/n^a} - (1/2\pi)(1 - (1/\lambda)^{-1/2}$$

$$\log \frac{\sqrt{1 + (1/\lambda)} + 1 - (1/n^a)}{\sqrt{1 + (1/\lambda)} - 1 + (1/n^a)}. \qquad \ldots(3.6)$$

Now since $(1 + 1/\lambda)^{1/2} = 1 + 1/2\lambda + O(1/\lambda^2)$ for all $|\lambda| < 1$ and also by noting that $a \to 1$, $\lambda \to \infty$ and $\lambda/n \to 0$ as $n \to \infty$, from (3.6) for all sufficiently large n we have

$$\int_0^{1-1/n^a} I_1(x)\, dx = (1/2\pi)(\log 2 + \log n)$$

$$- (1/2\pi)(1 + 1/\lambda)^{-1/2} \log \frac{2 + 1/2\lambda + O(1/\lambda^2) - 1/n}{1/2\lambda + O\{1/\lambda^2) + 1/n}$$

$$< (1/2\pi)(\log 2 + \log n) - (1/2\pi)(1 + 1/\lambda)^{-1/2}$$

$$\{\log \lambda + \log (4 - 1/n)\} < (1/2\pi) \log n/m^2 + 1. \qquad \ldots(3.7)$$

Also for $1 - 1/n^a\, 1 \leqq x \leqq 1$, since from Ibragimov and Maslova[4] $\Delta^{1/2}/\alpha < (2n - 1)^{1/2} (1 - x)^{-1/2}$, we have

$$\int_{1-1/na}^{1} I_1(x)\, dx < \int_{1-1/na}^{1} (2n - 1)^{1/2} (1 - x)^{-1/2}\, dx$$

$$= 2(2n^{1-a} - n^{-a})^{1/2} = O(\log n)^{1/2}. \qquad \ldots(3.8)$$

Now we find an upper limit for $\int_0^1 I_2(x)\, dx$. We have

$$mx\gamma\alpha^{-3/2} = mx(1 - x^2)^{-1/2} \{(n - 1) x^{2n+1} - n x^{2n-1} + x\} (1 - x^{2n})^{-3/2} \qquad \ldots(3.9)$$

which gives

$$| m (\gamma x - \alpha) \alpha^{-3/2} | = | m | (1-x^2)^{-1/2} (1-x^{2n})^{-3/2} | (2x^2 - 1)(1-x^{2n}) - n x^{2n} (1 - x^2) | . \quad \text{...(3.10)}$$

Now for $0 \leqq x \leqq \sqrt{3}/2$ given any ϵ positive exists an integer n_0 such that for all $n \geqq n_0$

$$n x^{2n} (1 - x^2) < n (3/4)^n < \epsilon.$$

We note that $(2x^2 - 1)(1 - x^{2n}) < (2x^2 - 1) < \epsilon$ only for a small interval of $\frac{1}{2}(1-\epsilon) < x^2 < \frac{1}{2}(1+\epsilon)$. Hence from (3.10) for all sufficiently large n we have

$$\int_{\sqrt{(1-\epsilon)/2}}^{\sqrt{(1+\epsilon)/2}} I_2 (x)\, dx < | m | (\sqrt{2/\pi}) \{1 - (3/4)^n\} \{\epsilon^2 + O(\epsilon^3)\}. \quad \text{...(3.11)}$$

Using the fact that ϵ can be made arbitrarily small we can see that the above integral tends to zero as $n \to \infty$ (for example choosing $\epsilon = 1/m$). On the other hand, let $\int_0^{1-1/n} \ldots dx$ indicate the integral over $0 < x \leqq 1 - 1/n$ excluding $\sqrt{(1-\epsilon/2)} \leqq x \leqq \sqrt{(1+\epsilon)/2}$, then from (3.10) and since for $0 \leqq x \leqq 1 - 1/n$ and n sufficiently large $x^{2n} < (1-1/n)^{2n} \to e^{-2}$, we have

$$\int_0^{1-1/n} I_2 (x)\, dx \leqq | m | (\sqrt{2/\pi}) (1 - e^{-2})^{-1/2} \int_0^{1-1/n} (1 - 2x^2)(1 - x^2)^{-1/2}$$

$$\exp\{- m^2 x^2 (1 - x^2)\} \leqq | m | (\sqrt{2/\pi}) (1 - e^{-2})^{-1/5} \int_0^{1-1/n}$$

$$(1- \; -2x^2)(1 -x^2)^{-1/2} \exp\{- m^2 x^2 (1 -x^2)\}\, dx$$

$$< (2\pi)^{-1/2} (1 - e^{-2})^{-1/2} \int_0^k u^{-1/2} e^{-u}\, du$$

where $u = m^2 x^2 (1 - x^2)$ and $k = m^2 (1 - 1/n)^2 (2/n - 1/n^2)$. Now by integrating by parts and since $u^{1/2} e^{-u} < 1$ for all sufficiently large n, we have

$$\int_0^{1-1/n} I_2 (x)\, dx < (2\pi)^{-1/2} (1 - e^{-2})^{-1/2} [\{2 (m^2/n)(2 - 1/n)(1 - 1/n)^2\}^{1/2}$$

$$\exp\{-(m^2/n)(2 - 1/n)(1 - 1/n)^2\} + 2 (m^2/n)(1 - 1/n)^2$$

$$(2 - 1/n)\}] \quad \text{...(3.12)}$$

which also tends to zero as $n \to \infty$. In order to estimate $\int_{1+1/n}^{1} I_2 (x)\, dx$ we note that

always $\gamma > \alpha/x$, then

$$| \gamma x/\alpha^{3/2} - 1/\alpha | < 2 | \gamma x/\alpha^{3/2} | . \quad \text{...(3.13)}$$

Since

$$\gamma + \sum_{i=1}^{n-1} i\, x^{2n-1} < (n/x) \sum_{i=1}^{n-1} x^{2i} = (n/x)\, \alpha$$

for $1 \leqq x \leqq 1 - 1/n$ we can obtain

$$\gamma x/\alpha^{3/2} \leqq n\, \alpha^{-1/2} \leqq \sqrt{n}\, (1 - 1/n)^{-(n-1)}.$$

This and (3.13) gives

$$\int_{1-1/n}^{1} I_2 (x)\, dx < (2 | m | /\pi) \int_{1-1/n}^{1} (\gamma x/\alpha^{3/2})\, dx$$

$$< (2 | m | /\pi \sqrt{n})\, (1 - 1/n)^{-(n-1)} \quad \text{...(3.14)}$$

which also tends to zero as $n \to \infty$. From (3.7), (3.8), (3.11), (3.12) and (3.14) we have

$$EN\,(0, 1) < (1/2\pi) \log (n/m^2) + O\, \{\log n\}^{1/2}. \quad \text{...(3.15)}$$

In order to obtain a lower estimate of $EN\,(0, 1)$ without loss of the generality we can assume $2m^2 > 1$. Then from (3.4) and since for all $0 \leqq x \leqq 1$

$$(1 - x^2)^2/2 - x^2 (1 - x^2) + x^2 (1 + x^2)/2 \leqq 1$$

we have

$$\int_{9}^{1-1/n^a} I_1 (x)\, dx > (1/\pi) \int_{0}^{1-1/n^a} (\Delta^{1/2}/\alpha) \exp \{-m^2 (1 - x^2)/(1 - x^{2n})\}\, dx$$

$$\geqq (1/2\pi) \int_{0}^{1-1/n^a} (1 - x)^{-1} \exp \{- 2m^2 (1 - x)\}\, dx$$

$$= (1/2\pi) \int_{2m^2/n^a}^{2m^2} u^{-1} e^{-u}\, du = (1/2\pi) \{ \int_{2m^2/n^a}^{2m^2} u^{-1}\, du$$

$$- \int_{2m^2/n^u}^{1} u^{-1} (1 - e^{-u})\, du - \int_{1}^{2m^2} u^{-1}\, du + \int_{1}^{2m^2} u^{-1} e^{-u}\, du\} \quad \text{...(3.16)}$$

where $u = 2m^2 (1 - x)$. Now by noticing that $u^{-1} (1 - e^{-u}) < 1$ and $\int_{1}^{2m^2} u^{-1} e^{-u}\, du$ is always positive, from (3.19) for all sufficiently large n we have

$$\int_{0}^{1-1/n^a} I_1 (x)\, dx > (1/2\pi) \{a \log n - \log m^2 - 1 - \log 2\}$$

$$= (1/2\pi) \log (n/m^2) + O\, (\log \log n). \quad \text{...(3.17)}$$

Finally from (3.15) and (3.17) we obtain the asymptotic formula

$$EN\,(0,\,1) \sim \log\,(n/m^2).$$

On the other hand for m bounded from (3.5) and (3.8) we have

$$\int_0^1 I_1\,(x)\,dx < (1/\pi) \int_0^{1-1/n^a} (1-x^2)^{-1}\,dx + (1/\pi) \int_{1-1/n^a}^{1} (2n-1)^{1/2}\,(1-x)^{-1/2}$$

$$\times\,dx < (1/2\pi)\log n + O\,\{\log n\}^{1/2}.$$

From this and (3.11), (3.12) and (3.14) we have

$$EN\,(0,\,1) < (1/2\pi)\log n + O\,\{\log n\}^{1/2}. \qquad \ldots(3.18)$$

Also in order to obtain a lower estimate, from (3.16) for all sufficiently large n we have

$$\int_0^{1-1/n^a} I_1\,(x)\,dx > (a/2\pi)\log n - 2m^2$$

$$= (1/2\pi)\log n + O\,\{\log\log n\}. \qquad \ldots(3.19)$$

Hence from (3.18) and (3.19) for m bounded we have

$$EN\,(0,\,1) \sim (1/2\pi)\log n.$$

Now we shall find the asymptotic relation for $EN\,(1,\,\infty)$. Let $y = 1/x$ then we have

$$\int_1^{\infty} I\,(x)\,dx = \int_0^1 I\,(1/y)\,y^{-2}\,dy. \qquad \ldots(3.20)$$

Since from Farahmand[2] (p. 707) we have $\Delta\,(1/y) = y^{-2(2n-4)}\,\Delta\,(y)$ and $\alpha\,(1/y) = y^{-(2n-2)}$ $\alpha\,(y)$, from (2.2) and (3.20) we have

$$\int_1^{\infty} I_1\,(x)\,dx < (1/\pi)\int_0^1 y^2\,\{\Delta\,(1/y)\}^{1/2}/\alpha(1/y)\}\,dy$$

$$= (1/\pi)\int_0^1 \{\Delta\,(y)\}^{1/2}/\alpha\,(y)\,dy$$

$$< (1/\pi)\int_0^{1-1/n} (1-y^2)^{-1}\,dy + (1/\pi)\int_{1-1/n}^{1} (2n-1)^{1/2}\,(1-y)^{-1/2}\,dy$$

$$< (1/2\pi)\log n + 3.5. \qquad \ldots(3.21)$$

Now we find an upper estimate for $I_2\,(x)$. We have

$$\{(m/y)\,\gamma\,(1/y) - m\,(\alpha\,(1/y)\}\,\{\alpha\,(1/y)\}^{-3/2}$$

$$= m\,y^{n-1}\,(1-y^2)^{-1/2}\,(1-y^{2n})^{-3/2}\,\{n\,(1-y^2) - (2-y^2)\,(1-y^{2n})\}. \qquad \ldots(3.22)$$

Let a be the same constant as before. From (2.2), (3.13) and (3.22) and for all sufficiently large n we have

$$\int_0^{1-1/n^a} y^{-2} I_2(1/y)\, dy \leqq (2\,|\,m\,|\,n/\pi) \int_0^{1-1/n^a} y^{n-3}(1-y^2)^{1/2}(1-y^{2n})^{-3/2}\, dy$$

$$\leqq (2\,|\,m\,|\,n/\pi)\exp(-n^{1-a})\{1-\exp(2n^{1-a})\}/(n-2) \qquad \ldots(3.23)$$

which, since $\exp(-n^{1-a}) = n^{-10}$, tends to zero as $n \to \infty$. On the other hand for $1-(1/n^a) \leqq y \leqq 1$

$$x\gamma(x)\{\alpha(x)\}^{-3/2} < n\{\alpha(x)\}^{-1/2} = n\,y^{n-1}(1-y^2)^{1/2}(1-y^{2n})^{-1/2}$$

$$< n\,y^{n-1}\{2n^{-a}-n^{-2a}\}^{1/2}\{1-(1-(1/n^a)^{2n}\}^{-1/2}.$$

Hence from this, (2.2) and (3.13) we have

$$\int_{1-1/n^a}^{1} y^{-2} I_2(1/y)\, dy \leqq (2\,|\,m\,|\,n\sqrt{2/\pi})\,n^{-a/2}\{1-(1/n^a)^{2n}\}^{-1/2}/n-2$$

$$= O\{\log n\}^{1/2}. \qquad \ldots(3.24)$$

Since $n^{-a/2} = (n^{-1}\log n^{10})^{1/2}$. Hence from (3.21), (3.23) and (3.24) we obtain

$$EN(1,\infty) < (1/2\pi)\log n + O\{\log n\}^{1/2}. \qquad \ldots(3.25)$$

To get the lower estimate for $EN(1,\infty)$, for $0 \leqq y \leqq 1-1/n^a$ we have

$$\alpha = y^{-(2n-2)}(1-y^{2n})/(1-y^2),\ \gamma = y^{-(2n-3)}\{1-(1-y^{2n})/(1-y^2)\}/(1-y^2),$$

$$\Delta = y^{-(4n-8)}\{(1-y^{2n})^2(1-y^2)^{-4} - n^2y^{2n-2}(1-y^2)^{-2}\}$$

$$= y^{-(4n-8)}(1-y^{2n})^2(1-y^2)^{-4} + O\{n^{2+2a}\exp(-2n^{1-a})\}$$

and

$$\beta/\Delta = y^{2n-4}(1-y^2)(1-y^{2n})^{-1}\{1+y^2+n^2(1-y^2)^2(1-y^{2n})^{-1}$$
$$-2n(1-y^2)(1-y^{2n})\} + O\{n^{2+2a}\exp(-2n^{1-a})\},$$

Hence, from these, and simple algebra, we find

$$(-\alpha + 2\gamma/y - \beta/y^2)/\Delta = y^{2n-6}(1-y^2)(1-y^{2n})^{-1}[(1-y^2)^2$$
$$-2(1-y^2)^2(1-y^{2n})^{-1}\{1-(1-y^{2n})/(1-y^2\} + 1 + y^2 + n^2$$
$$(1-y^2)^2/(1-y^{2n}) - 2n(1-y^2)/(1-y^{2n})]$$
$$+ O\{n^{2+2a}\exp(-2n^{1-a})\} < n^2\exp(2n^{1-a}) = n^{-18} \qquad \ldots(3.26)$$

for all sufficiently large n. Now, since from (3.26), $\exp\{m^2(-\alpha + 2\gamma/y - \beta/y^2)/2\Delta\}$ tends to one as $n \to \infty$, from (3.20), (3.26) and for all sufficiently large n

$$EN(1, \infty) > \int_1^\infty I_1(x)\,dx > (1/\pi) \int_0^{1-1/n^a} (1 - y^2)^{-1}\,dy$$

$$= (a/2\pi) \log n - (1/2\pi) \log (2 - 1/n^a) = (1/2\pi) \log n + O(\log \log n). \qquad \ldots(3.27)$$

Finally from (3.25) and (3.27) we have asymptotic formula

$$EN(1, \infty) \sim (1/2\pi) \log n.$$

Sinnce a_j and $-a_j$ ($j = 0, 1, \ldots, n-1$) both have the standard normal distribution

$EN(-1, 0) = EN(0, 1)$ and $EN(-\infty, -1) = EN(1, \infty)$, we have proof of the theorem.

If we add this result to asymptotic formula in Farahmand[2] for k constant such that (k^2/n) tends to zero we can obtain the result for the case in which $P(x) = mx + k$. For this case

$$\omega = -mx - k \text{ and } \eta = \{\gamma(mx + k) - \alpha m\}/\alpha^{1/2} \Delta^{1/2}$$

while α, β, γ and μ will remain the same. The Kac-Rice formula will be

$$EN(a, b) = \int_a^b (\Delta^{1/2}/\pi\alpha) [\exp [\{-\alpha m^2 + 2m\gamma (mx + k) - \beta (mx+k)^2\}/2\Delta]$$

$$+ [\{-\alpha m + \gamma (mx + k)\}/\pi\alpha^{3/2}] \exp \{-(mx + k)^2/2\alpha\}$$

$$\times \operatorname{erf} [\{-\alpha m + \gamma (mx + k)\}/\alpha^{1/2} \Delta^{1/2}]\}\,dx$$

and it is interesting to know for this case

$EN(-1, 1) \sim (1/\pi) \log (n/(m^2 + k^2)\}$ if either m or k or both tend to infinity as $n \to \infty$

$EN(-1, 1) \sim (1/\pi) \log n$ if m and k are bounded

$EN(-\infty, 1) = EN(1, \infty) \sim (1/2\pi) \log n.$

References

1. H. Cramer and M. R. Leadbetter, *Stationary and related Stochastic Process*. Wiley, New York 1967.
2. K. Farahmand, *Annals Probability* 14 (1986), 702–709.
3. I. A. Ibragimov and N. B. Maslova, *Theor. Probab. Appl.* 16 (1971), 228–48.
4. I. A. Ibragimov and N. B. Maslova, *Soviet Math. Dokl.* 12 (1971), 1004–1008.
5. M. Kac, *Bull. Am. Math. Soc.* 49 (1943), 314–20.
6. B. F. Logan, L. A. Shepp, *Proc. London Math. Soc.* (8) 18 (1968), 308–14.
7. M. Sambandham, *Indian J. pure appl. Math.* 7 (1976), 1062–70.
8. D. C. Stevens, *Commnn. pure appl. Math.* 22 (1969), 457–77.
9. S. O. Rice, *Bell. System Tech. J.* 25 (1945), 46–156.

Indian J. pure appl. Math., 20(1) : 10–39, January 1989

GENERALIZED CONVEXITY IN MULTI-OBJECTIVE PROGRAMMING

M. N. VARTAK* AND INDRANI GUPTA

Department of Mathematics, Indian Institute of Technology, Powai, Bombay 400076

(*Received* 17 *March* 1986; *after revision* 6 *June* 1988)

In this paper the relationships between solutions of (1) the weighting problem, (2) the kth objective Lagrangian problem, (3) the kth objective ϵ-constraint problem associated with a multi-objective nonlinear programming problem (MONLP) and noninferior solutions of the MONLP problem itself are studied under weaker types of convexities such as quasiconvexity, pseudoconvexity and η-convexity. The same ideas are applied to study refinement of duality theorems for MONLP. In particular the role played by linear approximation problem associated with MONLP has been discussed in detail. The concept of selective duality has also been investigated under weaker conditions of convexity.

1. INTRODUCTION

We consider the following multi-objective programming problem, also known as vector optimization problem (VOP) :

Given $f : R^N \to R^n$, $g : R^N \to R^m$, $S \subseteq R^N$, find an $x^* \in X$ (if such an x^* exists) such that

$$f(x^*) = \min_{x \in X} f(x) \qquad \text{...(1.1.1)}$$

where

$$X = \{x \in R^N : g(x) \leqq 0, x \in S\}. \qquad \text{...(1.1.2)}$$

The following conventions for equalities and inequalities for vectors $x, y \in R^N$ are used (Mangasarian[12], p. 14):

$$x = y \Leftrightarrow x_i = y_i \text{ for all } i = 1, 2, \ldots, N;$$

$$x \geqq y \Leftrightarrow x_i \geqq y_i \text{ for all } i = 1, 2, \ldots, N;$$

$$x \geqslant y \Leftrightarrow x \geqq y \text{ but } x \neq y;$$

$$x > y \Leftrightarrow x_i > y_i \text{ for all } i = 1, 2, \ldots, N.$$

---

*Present address : Visiting Professor, NITIE, Vihar Lake, Bombay, 400087.

A vector $x^* \in X$ is said to be a noninferior (or an efficient or a nondominated or a Pareto-Optimal) solution of the VOP if there exists no other feasible solution $x \in X$ such that $f(x) \leqslant f(x^*)$. We denote the set of all noninferior solutions of the VOP by X^*.

The most common strategy of getting the set X^* is to obtain nondominated solutions of the VOP in terms of optimal solutions of some appropriately associated scalar optimization problems. The following are the three common approaches of obtaining a scalar problem to be associated with VOP for this purpose.

(i) *The Weighting Problem*

Let $W = \{w \in R^n; w \geqq 0, \sum_{j=1}^{n} w_j = 1\}$ be the set of nonnegative weights in R^n. For a given $w \in W$ the weighting problem $P(w)$ is defined as

$$P(w): \min_{x \in X} \sum_{j=1}^{n} w_j f_j(x).$$

(ii) *The* k*th-objective Lagrangian Problem*

For a given k such that $1 \leqq k \leqq n$, the kth-objective Lagrangian problem is defined as

$$P_k(u): \min_{x \in X} \{f_k(x) + \sum_{\substack{j=1 \\ j \neq k}}^{n} u_j f_j(x)\}$$

where

$$u \in U_{(k)} = \{(u_1, u_2, \ldots, u_{k-1}, u_{k+1} \ldots, u_n)^T : u_j \geqq 0,$$
$$j = 1, 2, \ldots, n; j \neq k\} \subseteq R^{n-1}.$$

(iii) *The* k*th-objective* ϵ*-constraint Problem*

For a given k such that $1 \leqq k \leqq n$ and a given $\epsilon_{(k)} = (\epsilon_1, \ldots, \epsilon_{k-1}, \epsilon_{k+1}, \ldots, \epsilon_n)^T \in R^{n-1}$, the kth-objective ϵ constraint problem is defined as

$$P_k(\epsilon_{(k)}): \min_{x \in X} f_k(x)$$

subject to $f_j(x) \leqq \epsilon_j$, $j = 1, 2, \ldots, n$; $j \neq k$. For a given point $x^0 \in R^N$, the problem $P_k(\epsilon^0_{(k)})$ is defined as

$$P_k(\epsilon^0_{(k)}): \min_{x \in X} f_k(x)$$

subject to $f_j(x) \leqq \epsilon^0_j = f_j(x^0)$, $j = 1, 2, \ldots, n;\ j \neq k$.

That is $P_k\,(\epsilon^0_{(k)})$ is exactly $P_k\,(f_{(k)}\,(x^0))$.

Chankong and Haimes[3] (p. 119) have given some fundamental results concerning characterization of a noninferior solution of the VOP in terms of optimal solutions of these three scalarized optimization problems with the help of an implication diagram which involves the concept of convex functions. Hanson[5] and Kaul and Kaur[7] have defined new concepts of η-convex, η-strictly convex, η-pseudoconvex and η-quasiconvex functions which are weaker generalizations of the corresponding concepts of convex, strictly convex, pseudoconvex and quasiconvex functions.

In the second section of this paper we derive noninferior solutions of the VOP in terms of the three scalar optimization problems (SOP's) under weaker convexity assumptions. In the third section we study duality theory for the VOP under some weaker convexity assumptions, the pair of primal-dual multi-objective programs being that mentioned by Bitran[2].

2. Vector Optimization Theory

2.1. *Relationships Among the Three Forms of Scalarization*

First we consider the scalar nonlinear problem P : Given $f:\ R^N \to R, g:\ R^N \to R^m$, $S \subseteq R^N$, find an $x^* \in X$ (if such an x^* exists) such that

$$f\,(x^*) = \min_{x \in X} f(x) \qquad \text{...(2.1.1)}$$

where

$$X = \{x \in R^N : x \in S, g\,(x) \leqq 0\}. \qquad \text{...(2.1.2)}$$

The standard Lagrangian function $L\,(x, v)$ of the problem P is defined by

$$L\,(x, v) = f\,(x) + v^T\,g\,(x)$$

where $v \in R^m$.

Let S be an open set in R^N. Let $\bar{x}$ be a local minimum of P (cf. Mangasarian[12], p. 93). If g satisfies any of the constraint qualifications (for example, the Kunh-Tucker constraint qualification or the Weak Arrow-Hurwicz-Uzawa constraint qualification (see Mangasarian[12])), then there exists a $\bar{v} \in R^m$ such that $(\bar{x}, \bar{v})$ satisfies the Kuhn-Tucker conditions :

(a) $\nabla_{\bar{x}}\ L\,(\bar{x}, \bar{v}) = 0$

(b) $g\,(\bar{x}) \leqq 0$

(c) $\bar{v}^T\,g\,(\bar{x}) = 0$

(d) $\bar{v} \geqq 0$.

...(2.1.3)

When $x \in R^N$ and $u(x) \in R^n$, we shall use the symbol $\nabla_x u(x)$ to denote the $N \times n$ matrix, the jth column of which is the gradient of the component $u_j(x)$ of $u(x)$ w.r.t. x. Similar meaning is given to $\nabla_{\bar{x}} L(\bar{x}, \bar{v})$. For the sake of brevity, we shall also write $\nabla u(x)$ for $\nabla_x u(x)$.

Let I be the set of indices of active constraint functions at $\bar{x}$, i. e. $I = \{i : g_i(\bar{x}) = 0\}$. Let $I' = \{i : g_i(\bar{x}) < 0\}$. Let g_J. and $g_{I'}$ be the subvectors of g corresponding to the index sets I and I', respectively. Let $v \in R^m$ be such that $v \geqq 0$ and $v^T g(\bar{x}) = 0$. Clearly $v_I \geqq 0$ and $V_I' = 0$ where v_I and $v_{I'}$ denote the subvectors of v corresponding to the index sets I and I', respectively. Let $K = \{i \in I : v_i = 0\}$ and $J = \{i \in I : v_i > 0\}$, and let g_J and g_K be the subvectors of g_I corresponding to the index sets J and K, respectively. The subvectors v_J and v_K of v_I are similarly defined. Now let $J1$, $J2$, $J3$ be any subsets of J such that $J1 \cup J2 \cup J3 = J$, $J1 \cap J2 = J2 \cap J3 = J3 \cap J1 = \phi$. Finally let v_{J_1}, v_{J_2}, v_{J3} and g_{J_1}, g_{J_2}, v_{J3} be the subvectors of v_J and g_J, respectively, corresponding to the sets $J1$, $J2$, $J3$.

The $(N + m)$ vector $(\bar{x}, \bar{v})$ which satisfies the Kuhn-Tucker condition (2.1.3), is said to satisfy the McCormick[13] second order sufficient optimality condition if f and g are twice differentiable and

$$\left.\begin{array}{l} d \in R^N \\ \nabla g_J(\bar{x})^T d = 0 \\ \nabla g_K(\bar{x})^T d \leqq 0. \end{array}\right\} \Rightarrow d^T \nabla_{\bar{x}\bar{x}} L(\bar{x}, \bar{v})\, d > 0. \qquad \ldots(2.1.4)$$

The following theorem has been proved by Mahajan[11] in his unpublished Ph.D. thesis.

Theorem 2.1.1 [Mahajan[11], Theorem 1.4.2 p. 37]—Consider the scalar problem P defined by (2.1.1), (2.1.2). Let f and g be differentiable at $\bar{x} \in X$. If there exists a $v \in R^m$ such that $(\bar{x}, v)$ satisfies

(i) $(\nabla f(\bar{x}) + \nabla g(\bar{x})\, v)^T (x - \bar{x}) \geqq 0$ for all $x \in X$

(ii) $v \geqq 0$

(iii) $v^T g(\bar{x}) = 0$

and if, for any arbitrary disjoint subsets $J1$, $J2$, $J3$ of J whose union is J,

(iv) $f + v_{J1}^T g_{J_1}$ is pseudoconvex at $\bar{x}$ w.r.t. X,

(v) $v_{J2}^T g_{J_2}$ is differentiable and quasiconvex at $\bar{x}$ w. r. t. X, and

(vi) g_{J3} is differentiable and quasiconvex at $\bar{x}$ w. r. t. X, then $\bar{x}$ is a global optimal solution of P.

Mahajan[11] has also pointed out three particular cases of the above result corresponding to

$$J1 = J,\ J2 = J3 = \phi \qquad \text{...(2.1.5a)}$$

$$J2 = J,\ J1 = J3, = \phi \qquad \text{...(2.1.5b)}$$

$$J3 = J,\ J1 = J2 = \phi. \qquad \text{...(2.1.5c)}$$

(See Mahajan[11], Remark 1 4.2, p. 39).

Define

$$f^T_{(k)} = (f_1, f_2, \ldots, f_{k-1}, f_{k+1}, \ldots, f_n)^T.$$

Consider the problem $P_k\,(f_{(k)}\,(x^0))$ for a given point $x^0 \in X$. Under a suitable constraint qualification (for example the Kuhn-Tucker or the weak Arrow-Hurwicz-Uzawa constraint qualification given in Mangasarian[12]), the following Kuhn-Tucker conditions of optimality hold at a minimum x^* of $P_k\,(f_{(k)}\,(x^0))$.

$$\nabla f_k(x^*) + \nabla f_{(k)}(x^*)\, u + \nabla g(x^*)\, v = 0,\ u \in U_{(k)} \qquad \text{...(2.1.6a)}$$

$$v^T g(x^*) = 0 \qquad \text{...(2.1.6b)}$$

$$v_i \geqq 0,\ i = 1, 2, \ldots, m,\ u_j \geqq 0,\ j = 1, 2, \ldots, k-1, k, k+1, \ldots, n, \qquad \text{...(2.1.6c)}$$

$$g(x^*) \leqq 0.$$

Now we give the relations among the solutions of scalar optimization problems.

The following Lemma generalizes Lemma 4.6 of Chankong and Haimes[3] (p. 127) in the sense that convexity assumption is weakened.

Lemma 2.1.1—Let S be an open set in R^N. For a given $x^0 \in X$ let $x^* \in X$ solve $P_k\ (\epsilon^0_{(k)})$. Let f and g be differentiable at x^*. Assume that there exist $u \in U_{(k)}$ and a $v \in R^m$ such that (2.1.6) holds at (x^*, u, v).

Further assume that

(i) $f_k + u^T f_{(k)}$ is pseudoconvex at x^* w.r.t. X and g_J is quasiconvex at x^* w.r.t. X,

or

(ii) $f_k + u^T f_{(k)}$ is pseudoconvex at x^* w.r.t. X and $v^T_J\ g_J$ is quasiconvex at x^* w.r.t. X,

or

(iii) the numerical function $f_k + u^T f_{(k)} + v^T_J\ g_J$ is pseudoconvex at x^* w.r.t. X.

Then x^* solves $P_k(u)$.

PROOF : The lemma is an immediate consequence of the three cases of (2.1.5).

The proof of the following Lemma is trivial and is therefore omitted.

Lemma 2.1.2—Let S be an open set in R^N. For a given $x^0 \in X$ let x^* solve $P_k(\epsilon^0_{(k)})$. Let f and g be twice continuously differentiable at x^*. Let the Kuhn-Tucker constraint qualification (KTCQ) or the Weak-Arrow-Hurwicz-Uzawa constraint Qualification (WAHUCQ) hold at x^*. Let $U^0_{(k)} \subseteq U_{(k)}$ be such that for each $u \in U^0_{(k)}$ there exists a $v \in R^m$ such that (u, v) is a Kuhn-Tucker multiplier under the above C. Q. for $P_k(\epsilon^0_{(k)})$ and such that the following second order sufficient optimality condition holds for $P_k(u)$ at (u, v):

$$\left.\begin{array}{l} d \in R^N \\ \nabla g_J(x^*)^T d = 0 \\ \nabla g_k(x^*)^T d \leqq 0 \end{array}\right\} \Rightarrow d^T \nabla_{x^*x^*} L(x^*, u, v)\, d > 0 \qquad \ldots(2.1.7)$$

where $L(x^*, u, v) = f(x^*) + u^T f_{(k)}(x^*) + v^T g(x^*)$. Then x^* solves $P_k(u)$ for all $u \in U^0_{(k)}$.

Remark 2.1.1: Here we have imposed the second order sufficient optimality conditions in absence of convexity condition. The second order sufficient optimality condition automatically hold when $\nabla_{x^*x^*} L(x^*, u, v)$ is positive definite. Following is an example where convexity assumption does not hold for f but $\nabla_{x^*x^*} L(x^*, u, v)$ is positive definite.

minimize $(x^3, - x^3 + 2x, - 4x)$

subject to $x^6 - 1 \leqq 0$.

One of the associated constraint problems for $x^0 = 0.8$ is

$P_1(\epsilon^0_{(1)})$: minimize x^3

$$\begin{aligned} \text{subject to } - x^3 + 2x &\leqq 1.08 \\ - 4x &\leqq 3.2 \\ x^6 - 1 &\leqq 1. \end{aligned}$$

We find that $0.8 \leqq x \leqq 1$ are the feasible solutions and $x = 0.8$ solves $P_k(\epsilon^0_{(1)})$. $(0.8, u, v)$ where $u^T = (u_1, u_2)$, $u_1 \geqq 0$, $u_2 = \dfrac{1.92 + 0.08\, u_1}{4}$, $v = 0$, satisfies the Kuhn-Tucker conditions for $P_k(\epsilon^0_{(1)})$.

$$U^0_{(k)} = ((u_1, u_2) : 0 \leqq u_1 \leqq 1, u = \frac{1.92 + 0.08\, u_1}{4}).$$

$u \in U^0(k)$ and $v = 0$ are such that the second order sufficient optimality condition hold for $P_k(u)$, through f_1 is not convex. Hence 0.8 solves $P_k(u)$ for all $u \in U^0\ (k)$.

We now define η-convex, η-strictly convex, η-pseudoconvex and η-quasiconvex functions[5,7].

Definition 2.1.1—Let f be a numerical function defined on an open set $S \subseteq R^N$ and let f be differentiable at $x^0 \in S$. Then f is said to be

(i) η-convex at $x^0 \in S$ is $\exists$ a function $\eta : S \times S \rightarrow R^N$ such that, for all $x \in S$,

$$f(x) - f(x^0) \geqq \eta^T(x, x^0) \nabla f(x^0) \qquad ...(2.1.8)$$

(ii) η-strictly convex at $x^0 \in S$ if $\exists$ a function $\eta : S \times S \rightarrow R^N$ such that strict inequality holds in (2.1.8) for all $x \in S$, $x \neq x^0$,

(iii) η-pseudoconvex at $x^0 \in S$ if $\exists$ a function $\eta : S \times S \rightarrow R^N$ such that

$$\eta^T(x, x^0) \nabla f(x^0) \geqq 0 \Rightarrow f(x) \geqq f(x^0) \text{ for all } x \in S \qquad ...(2.1.9)$$

(iv) η-quasiconvex at $x^0 \in S$ if $\exists$ a function $\eta : S \times S \rightarrow R^N$ such that

$$f(x) \leqq f(x^0) \Rightarrow \eta^T(x, x^0) \nabla f(x^0) \leqq 0 \text{ for all } x \in S. \qquad ...(2.1.10)$$

In addition we introduce the definition of η-strictly pseudoconvex function.

Definition 2.1.2—Let f be a numerical function on an open set $S \subseteq R^N$ and let f be differentiable at $x^0 \in S$. Then f is said to be η-strictly pseudoconvex at $x^0 \in S$ if $\exists$ a function $\eta : S \times S \rightarrow R^N$ such that for all $x \in S$, $x \neq x^0$

$$\eta^T(x, x^0) \nabla f(x^0) \geqq 0 \Rightarrow f(x) > f(x^0). \qquad ...(2.1.11)$$

A function is said to have any of the above properties on S if it has that property at all points of S. By taking $\eta(x_1, x_2) = (x_1 - x_2)$ and by using the definition of Ponstein[14] it can be seen that every strictly pseudoconvex function is η-strictly pseudoconvex. The converse, however, is not true as can be seen from the following example (cf. Kaul and Kaur[7]).

Example —Let the function $f : S \rightarrow R$ be defined by $f(x) = x_1 + \sin x_2$. Let $S \subseteq R^2$ be given by

$$S = \{x \in R^2 : 4x_1^2 + 4x_2^2 - 9 \leqq 0,\ x_1, x_2 \geqq 0\}$$

and let the function $\eta : S \times S \rightarrow R^2$ be given by

$$\eta(x, u) = \begin{pmatrix} \dfrac{\sin x_1 - \sin u_1}{\cos u_1} \\ \dfrac{\sin x_2 - \sin u_2}{\cos u_2} \end{pmatrix}.$$

Then the function $f: S \to R$ defined by $f(x) = x_1 + \sin x_2$ is η-strictly pseudoconvex on S, but it is not strictly pseudoconvex on S since at $x = (\frac{\pi}{3}, 0)$, $u = (\frac{\pi}{6}, \frac{\pi}{3})$ we have $(x - u)^T \nabla f(u) = 0$ whereas $f(x) < f(u)$. Kaul and Kaur[7] have used the function in the above example to show that exery η.pseudoconvex function need not be pseudoconvex.

We define a function to be η-concave, η-strictly concave, η-pseudoconcave, η-strictly pseudoconcave, η-quasiconcave, at $x^0 \in S$ (on S) if and only if—f is η-convex, η-strictly convex, η-pseudoconvex, η-strictly pseudoconvex, η-quasiconvex at $x^0 \in S$ (on S) respectively.

The following Lemma is similar to Theorem 2.1.1.

Lemma 2.1.3—Let $\bar{x} \in S$. Assume that there evists a $\bar{v} \in R^m$ such that $(\bar{x}, \bar{v})$ safisfies the Kuhn-Tucker conditions (2.1.3). Further assume that

(i) f is η-pseudoconvex at $\bar{x}$ and g_J is η-quasiconvex at $\bar{x}$ for the same function η.

or

(ii) f is η-pseudoconvex at $\bar{x}$ and $\bar{v}_J^T\, g_J$ is η-quasiconvex at $\bar{x}$ for the same function η

or

(iii) $f + \bar{v}_J^T\, g_J$ is η-pseudoconvex at $\bar{x}$.

Then $\bar{x}$ is an optimai solution of problem P defined by (2.1.1) and (2.1.2).

The Lemma stated below is similar to Lemma 2.1.1 above.

Lemma 2.1.4—Let x^* solve $P_k\ (\epsilon^0_{(k)})$ and let (x^*, u, v), $u \in U_{(k)}$ $v \in R^m$ satisfy the Kuhn-Tucker conditions (2.1.6) of optimality. Let

(i) $f_k + u^T f_{(k)}$ be η-pseudoconvex at x^* and g_J be η-quasiconvex at x^* for the same function η,

or

(ii) $f_k + u^T f_{(k)}$ be η-pseudoconvex at x^* and $v_J^T\, g_J$ be η-quasiconvex at x^* for the same function,

or

(iii) $f_k + u^T f_{(k)} + v_J^T\, g_J$ be η-pseudoconvex at x^*. Then x^* solves $P_k\,(u)$.

Let (x^*, u, v), $u \in U_k$, $v \in R^m$ satisfy the Kuhn-Tucker conditions (2.1.6) of optimality for $P_k\,(\epsilon^0_{(k)})$. Then (2.1.6a) can be written in the form

$$\sum_{j=1}^{n} u'_j \nabla f_j(x^*) + \sum_{j=1}^{m} v_l \nabla g_l(x^*) = 0$$

where

$$u'_j = u_j \text{ for } j \neq k, u'_k = 1.$$

Dividing throughout by $\sum_{j=1}^{n} u'_j$ and writing

$$w_j = u'_j \left(\sum_{j=1}^{n} u'_j \right)^{-1} \text{ and } \bar{v}_l = \bar{v}_l \left(\sum_{j=1}^{n} u'_j \right)^{-1}$$

(2.1.6a) can now be written as

$$\sum_{j=1}^{n} w_j \nabla f_j(x^*) + \sum_{j=1}^{m} \bar{v}_l \nabla g_l(x^*) = 0. \quad \text{...(2.1.12a)}$$

Moreover, we note that $w_j \geqq 0$ and $\sum_{j=1}^{n} w_j = 1$.

Similarly (2.1.6b) through (2.1.6d) can be written as

$$\bar{v}_i^T g_l(x^*) = 0 \quad i = 1, 2, \ldots, m, \quad \text{...(2.1.12b)}$$

$$\bar{v}_l \geqq 0 \quad i = 1, 2 \ldots, m, \quad \text{...(2.1.12c)}$$

$$g_l(x^*) \leqq 0 \quad i = 1, 2, \ldots, m. \quad \text{.. (2.1.12d)}$$

This leads to the following lemma which is a generalization of Lemma 4.1 of Chankong and Haimes[8].

Lemma 2.1.5—Let x^* solve $P_k(\epsilon^0_{(k)})$ and let $(x^*, \bar{v}, w)$, $\bar{v} \in R^m$, $w \in W$ satisfy (2.1.12). Assume that

(i) $w^T f$ is pseudoconvex (η-pseudoconvex) at $x^* \in S$ and g_J is quasiconvex (η-quasiconvex for the same function η) at x^*.

or

(ii) $w^T f$ is pseudoconvex (η-pseudoconvex) at $x^* \in S$ and $\bar{v}_J^T g_J$ is quasiconvex (η-quasiconvex for the same function η) at x^*

or

(iii) $w^T f + \bar{v}_J^T g_J$ is pseudoconvex or η-pseudoconvex at x^*. Then x^* solv es $P(w)$.

2.2. *Characterizing Noninferior Solutions in Terms of the Solutions of the* k*th-Objective ϵ-Constraint Problem.*

We now establish the relationship between the noninferior solutions of vector optimization problem and the kth-objective ϵ-constraint problem.

Theorem 2.2.1 Let S be an open set in R^N. Let x^* solve $P_k\ (\epsilon^*_{(k)})$. Let $f_{(k)}$ be η-quasiconvex at x^* and f_k be η-strictly pseudoconvex and η-quasiconcave at x^* for the same function η. Then $x^* \in X^*$.

PROOF : Since x^* solves $P_k\ (\epsilon^*_{(k)})$

$$f_k(x^*) = \min_{x \in X} f_k(x) \qquad \text{...(2.2.1)}$$

$$\text{subject to } f_{(k)}(x) \leqq f_{(k)}(x^*). \qquad \text{...(2.2.2)}$$

Since f_k is η-quasiconcave at x^* from (2.2.1)

$$\eta^T(x, x^*) \nabla f_k(x^*) \geqq 0, \text{ for all } x \in X \qquad \text{...(2.2.3)}$$

and $f_{(k)}$ is η-quasiconvex at x^* for the same function η,

$$\eta^T(x, x^*) \nabla f_{(k)}(x^*) \leqq 0.$$

Hence,

$$\left.\begin{array}{l} \eta^T(x, x^*) \nabla f_k(x^*) < 0 \\ \eta^T(x, x^*) \nabla f_{(k)}(x^*) \leqq 0 \end{array}\right] \text{has no solution } x \in X. \qquad \text{...(2.2.4)}$$

Suppose $x^* \notin X^*$. Then there exists an $\hat{x} \in X$, $\hat{x} \neq x^*$ such that $f(\hat{x}) \leqslant f(x^*)$ and since f_k is η-strictly pseudoconvex at x^* and $f_{(k)}$ is η-quasiconvex at x^*

$$\eta^T(\hat{x}, x^*) \nabla f_k(x^*) < 0$$

$$\eta^T(\hat{x}, x^*) \nabla f_{(k)}(x^*) \leqq 0$$

contradicting (2.2.4). Hence $x^* \in X^*$.

The following corollary is an immediate consequence.

Corollary 2.2.2—Let S be an open set in R^N. Let x^* solve $P_k\ (\epsilon^*_{(k)})$. Let $f_{(k)}$ be quasiconvex at x^* and f_k be strictly pseudoconvex and quasiconcave at x^*. Then $x^* \in X^*$.

The well known results regarding the noninferiority for vector optimization problem of solutions of the kth-objective constraint problem are the following (cf. Lin[9,10], Chankong and Haimes[3]) :

(i) If x^* solves $P_k(\epsilon^*_{(k)})$ for some k and if the solution is unique, then x^* is a noninferior solution of VOP.

(ii) If x^* solves $P_k(\epsilon^*_{(k)})$ for every $k = 1, 2, \ldots, n$ then $x^* \in X^*$.

(iii) Let x^* solve $P_k(\epsilon^*_{(k)})$ for some k. For this k, let Y_k be the set of all $\epsilon_{(k)} \in R^{n-1}$ such that $P_k(\epsilon_{(k)})$ is feasible. Let $\phi_k(\epsilon_{(k)}) = \inf\{f_k(x) : x \in X, f_j(x) \leqq \epsilon_j$ for each $j \neq k\}$ and let $\hat{Y}_k = \{\epsilon_{(k)} : \epsilon_{(k)} \in Y_k, \phi_k(\epsilon_{(k)}) > -\infty$ and there exists an $x^0 \in X$ such that $f_k(x^0) = \phi_k(\epsilon_{(k)})\}$. Then $x^* \in X^*$ iff $\phi_k(\epsilon_{(k)}) > \phi_k(\epsilon^*_{(k)})$ for all $\epsilon_{(k)} \in \hat{Y}_k$ such that $\epsilon_{(k)} \leqq \epsilon^*_{(k)}$.

To generate noninferior solutions using the results (i) to (iii) is not very practical. For instance, in (i), (ii) and (ii) the vector x^* which appears in the constrainrs of $P_k(\epsilon^*_{(k)}$ is assumed to be its solution which is rather difficult to comprehend. For a convex problem $P_k(.)$, strict convexity of the (primary) objective function $f_k(.)$ guarantees uniqueness without further checking. The above theorem shows that if the (primary) objective function $f_k(.)$ is strictly pseudoconvex as well as quasiconcave at the point x^* and the other objective functions are quasiconvex at x^* then it guarantees noninferiority.

As an example we see that the function $f(x) = -\cos^2(x)$, $x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ is both strictly pseudoconvex and quasiconcave at the point $x^* = 0$.

2.3. *Characterizing Noninferior Solutions in Terms of the Solutions of the Weighting Problem*

The following theorem is a generalization of Theorem 4.5 of Chankong and Haimes.[3].

Theorem 2.3.1—Let S be an open set in R^N. Let $x^* \in X^*$ and let $x^*, \bar{v}, w)$ satisfy (2.1.12). Assume that

(i) $w^T f$ is pseudoconvex (η-pseudoconvex) at x^* and g_J is quasiconvex (η-quasiconvex for the same function η) at x^*,

or

(ii) $w^T f$ is pseudoconvex (η-pseudoconvex) at x^* and $\bar{v}_J^T g_J$ is quasiconvex (η-quasiconvex for the same function η) at x^*,

or

(iii) $w^T f + \bar{v}_J^T g_J$ is pseudoconvex (η-pseudoconvex at x^*). Then x solves $P(w)$.*

PROOF : Since $x^* \in X^*$ it solves $P_k\,(\epsilon^*_{(k)})$ for all $k = 1, 2, \ldots, n$ (Chankong and Haimes[3], Theorem 4.1). Now the result follows from Lemma 2.1.5.

The following theorem gives the conditions under which a solution of the weighting problem is a noninferior solution of the vector optimization problem.

Theorem 2.3.2—Let S be an open set in R^N. Let x^* solve $P\,(w)$. Assume that g satisfies the KTCQ or WAHUCQ at x^*. Let $w_k \neq 0$ and the corresponding f_k be strongly quasiconvex (cf. Bazaraa and Shetty[1]) and pseudoconvex at x^*, $f_{(k)}$ and g be quasiconvex at x^*. Then $x^* \in X^*$.

PROOF : The vector g satisfies the KTCQ or WAHUCQ at $x^* \in X$; so there exists a $v \in R^m$ such that

$$\sum_{i=1}^{n} w_j\, \nabla f_j\,(x^*) + \sum_{i=1}^{m} v_i\, \nabla g_i\,(x^*) = 0;$$

$$\sum_{i=1}^{m} v_i\, g_i\,(x^*) = 0,$$

$$v_i \geqq 0$$

$$g_i\,(x^*) \leqq 0,\ i = 1, 2, \ldots, m,$$

or, since $w_k \neq 0$, for appropriate w'_j a dn $'_i$,

$$\nabla f_k\,(x^*) + \sum_{\substack{j=1 \\ j\neq 1}}^{n} w'_j\, \nabla f_j\,(x) + \sum_{i=1}^{m} v'_i\, \nabla g_i\,(x^*) = 0.$$

These are the Kuhn-Tucker conditions of optimality for $P_k\,(\epsilon^*_{(k)})$. We have f_k pseudoconvex at x^* and g and $f_{(k)}$ quasiconvex at x^*. Then by sufficient optimality theorem (Mangasarian[12], Theorem 10.1.1) x^* is a solution of $P_k\,(\epsilon^*_{(k)})$.

Also, the set $x^2 = \{x \in R^N : g\,(x) \leqq 0,\ f_{(k)}\,(x) - f_{(k)}\,(x^*) \leqq 0\}$, of feasible solution of $P_k\,(\epsilon^*_{(k)})$ is convex set in R^N and f_k is strongly quasiconvex at x^*. Hence x^* is the unique solution of $P_k\,(\epsilon^*_{(k)})$ (Bazarra and Shetty[1], Theorem 3.5.9). We now apply the well the known result (Theorem 4.2, p. 129, Chankong and Haimes[3]) that if x^* solves $P_k\,(\epsilon^*_{(k)})$ for some k and if the solution is unique, then x^* is noninferior solution of the VOP.

The well known conditions under which x^* is a non-inferior solution of the VOP when it solves $P(w)$ are (Chankong and Haimes[3], Theorem 4.6):

(i) $w_j > 0$ for all $j = 1, 2, \quad , n$

or

(ii) x^* is the unique solution of $P(w)$.

In the above theorem we have shown the noninferiority with some weaker convexity assumption without assuming uniqueness condition or the condition that all the weights are positive.

It has been shown that a function is strictly pseudoconvex if it is strongly quasiconvex and Pseudoconvex (Ponstein[14]). There are, however, functions which are strongly quasiconvex but not pseudoconvex and vice versa. For example, the function $f(x) = x + x^3, x \in R$ is strictly pseudoconvex every-where. The function $f(x) = x^3$ is strongly quasiconvex at $x = 0$ but is not pseudoconvex at $x = 0$. The function $f(x) = \sin x + \cos x, x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ is not strongly quasiconvex at $x = 0$, but is pseudoconvex at $x = 0$.

Theorem 2.3.3—Let S be an open set in R^N. Let x^* solve $P(w)$ for a given $w \in W$. Assume that $W^T f$ is η-quasiconcave at x^* and the function f is η-strictly pseudoconvex at x^* for the same function η. Then x^* is a noninferior solution of the VOP.

PROOF : Since x^* solves $P(w)$, $w^T f(x) \geqq w^T f(x^*)$ for all $x \in X$ and the function $w^T f$ is η-quasiconcave at x^* which implies

$$\eta^T (x, x^*) \nabla f(x^*) w \geqq 0 \text{ for all } x \in X. \qquad \ldots(2.3.!)$$

Suppose $x^* \notin X^*$. Then there exists an $\hat{x} \in X$, $\hat{x} \neq x^*$ such that $f(\hat{x}) \leqslant f(x^*)$, and since f is η-strictly pseudoconvex at x^*, this implies that

$$\eta^T (x, x^*) \nabla f(x^*) < 0$$

and since $w \in W$

$$\eta^T (x, x^*) \nabla f(x^*) w < 0.$$

This contradicte (2.3.1). Hence $x^* \in X^*$.

As an immediate consequence, we get :

Corollary 2.3.4—Let S be an open set in R^N. Let x^* solve $P(w)$ for a given $w \in W$. Assume that $w^T f$ is differentiable and quasiconcave at x^* and the function f is strictly pseudoconvex at x^*. Then $x^* \in X^*$.

Remark 2.3.1: Theorem 2.3.3 also holds, if instead of the assumption that the function f is η-strictly pseudoconvex at x^*, either of the following assumptions is made:

(i) The components of f corresponding to the non-zero components of w are η-strictly pseudoconvex at x^*.

(ii) $w^T f$ is η-strictly pseudoconvex at x^*.

Remark 2.3.2: For example consider the VOP

$$\text{minimize } f^T(x) = (f_1(x), f_2(x), f_3(x))$$
$$= (-\cos x, \sin^2 x, x^3 + \sin x)$$

where the domain of definition of all the function is

$$X = \{x \in R : -\frac{\pi}{4} \leqq x \leqq \frac{\pi}{4}\}.$$

Take $w^T = (\frac{1}{2}, \frac{1}{2}, 0)$. Then it is easy to check that the objective function of the associated weighting problem $w^T f(x) = \frac{1}{2}(\sin^2 x - \cos x)$ is quasiconcave at $x^* = 0$ and the objective function $f(x)$ of the VOP is strictly pseudoconvex at $x^* = 0$. Moreover x^* solves the weighting problem $P(w)$. Hence $0 \in X^*$.

The apparently anomalous situation arising due to the assumption that x^* is involved in the definition of $P_k(\epsilon^*_{(k)})$ and at the same time solves $P_k(\epsilon^*_{(k)})$ made in Theorem 2.2.1 and the results stated thereafter, can be overcome to some extent through the following result.

Result 1—Let x^* solve $P_k(\epsilon^0_{(k)})$ for some given $x^0 \in X$ and let $(x^*, \bar{v}, w)$, $\bar{v} \in R^m$, $w \in W$ satisfy the Khun-Tucker conditions (2.1.12). Let

(i) $w^T f$ be η-pseudoconvex at x^* and g_J be η-quasiconvex at x^* for the same function η,

or

(ii) $w^T f$ be η-pseudoconvex at x^* and $\bar{v}_J^T g_J$ be η-quasiconvex at x^* for the same function η,

or

(iii) $w^T f + \bar{v}_J^T g_J$ be η-pseudoconvex at x^*.

Further assume that

(1) f_k is η-strictly pseudoconvex at x^*, $f_{(k)}$ and g are η-quasiconvex at x^*, for the same function η,

or

(2) $w^T f$ is η-quasiconcave at x^* and f is η-strictly pseudoconvex at x^* for the same function η.

Then $x^* \in X^*$.

PROOF : From the first set of assumptions we get x^* solves $P(w)$ and from the second set of assumptions we get $x^* \in x^*$.

2.4. *Characterizing Noninferior Solutions in Terms of the Solutions of the Lagranian Problem*

The following theorem is generalization of Theorem 4.7 in Chankong and Haimes[3] in the sense that convexity assumptions are weakened.

Theorem 2.4.1—Let S be an open set in R^N. Let $x^* \in X^*$ and assume that there exist a $u \in U_{(k)}$ and a $v \in R^m$ such that (2.1.6) holds at (x^*, u, v). Further assume that

(i) $f_k + u^T f_{(k)}$ is pseudoconvex (η-pseudoconvex) at x^* and g_J is quasiconvex (η-quasiconvex for the same function η) at x^*,

or

(ii) $f_k + u^T f_{(k)}$ is pseudoconvex (η-pseudoconvex) at x^* and $v_J^T g_J$ is quasiconvex (η-quasiconvex for the same function η) at x^*,

or

(iii) $f_k + u^T f_{(k)} + v_J^T g_J$ is pseudoconvex or η–pseudoconvex at x^*,

or

(iv) the second order sufficient optimality conditions (2.1.7) hold good.

Then x^* solves $P_k(u)$.

PROOF : Since (x^*, u, v) satisfies (2.1.6), by applying the three cases of (2.1.5), and Lemma 2.1.2 and Lemma 2.1.4 we get this theorem.

We prove the following theorems in the reverse direction which are generalizations of Theorem 4.8 given in Chankong and Haimes[3].

Theorem 2.4.2—Let S be an open set in R^N. Let x^* solve $P_k(u)$ for some $u \in U_{(k)}$. Assume $f_{(k)}$ is η-quasiconvex at x^* and f_k is η-strictly pseudoconvex and η-quasiconcave at x^* for the same function η. Then $x^* \in X^*$.

PROOF : Since x^* solves $P_k(u)$ for some $u \in U_{(k)}$, x^* also solves P_k $(\epsilon^*_{(k)})$ (Chankong and Haimes[3], Lemma 4.7), and by Corollary 2.2.2, x^* is a noninferior solution of the VOP.

Theorem 2.4.3—Let S be an open set in R^N. Let $x^* \in X$ solves $P_k(u)$ for some $u \in U_{(k)}$. Let g satisfy the KTCQ or WAHUCQ at x^*. Let f_k be strictly pseudoconvex at x^*, $f_{(k)}$ and g be quasiconvex at x^*. Then $x^* \in X^*$.

PROOF : Since x^* solves $P_k(u)$, it solves $P(w)$ with

$$w_j = u_j (1 + \sum_{j=1}^{n} u_j)^{-1}, \; j = 1, 2, \ldots, n; \; j \neq k,$$

$$w_k = (1 + \sum_{j=1}^{n} u_j)^{-1}$$

Then from Theorem 2.3.2, we get the result.

Theorem 2.4.4—Let S be an open set in R^N. Let x^* solve $P_k(u)$ for some $u \in U_{(k)}$. Let

$$w_j = u_j (1 + \sum_{j=1}^{n} u_j)^{-1}, \; j = 1, 1, \ldots, n, \; j \neq k,$$

$$w_k = (1 + \sum_{j=1}^{n} u_j)^{-1}$$

Let $w^T f$ be quasiconcave (η-quasiconcave) at x^* and f be strictly pseudoconvex (η-strictly pseudoconvex for the same function η) at x^*. Then $x^* \in X^*$.

PROOF : Since x^* solves $P_k(u)$ it solves $P(w)$ with the given values of w_j, $j = 1, 2, \ldots, n$ and then the theorem follows from Theorem follows from Theorem 2.3.3.

Remark similar to Remark 2.3.1 for Theorem 2.3.3 also applies to Theorem 2.4. 4 above.

We now give a result similar to Result 1, Section 2.3.

Result 2—For a given $x^0 \in X$ let x^* solve $P_k(\epsilon^0_{(k)})$ and let (x^*, u, v) satisfy (2.1.6). Assume that

(i) $f_k + u^T f_{(k)}$ is η-pseudoconvex at x^* and g_J is η-quasiconvex at x^* for the same function η,

or

(ii) $f_k + u^T f_{(k)}$ is η-pseudoconvex at x^* and v_J^T is η-quasiconvex at x^* for the same function η,

or

(iii) $f_k + u^T f_{(k)} + v_J^T$ is η pseudoconvex at x^*.

Further assume that $f_{(k)}$ is η-quasiconvex at x^* and f_k is η-strictly pseudoconvex at x^* and η-quasiconcave at x^* for the same function η. Then $x^* \in X^*$.

PROOF : From the first set of assumptions we get x^* solves $P_k(u)$ and from the next assumption we get x^* is a noninferior solution of the VOP.

2.5. *Kuhn-Tucker Necessary and Sufficient Conditions for Noninferiority*

Denfiition—A feasible point x^* for VOP ((1.1.1), (1.1.2)) is said to satisfy the Kuhn-Tucker conditions for noninferiority (KTCN) for the VOP if

(i) all f and g are differentiable and $S \neq \phi$

and

(ii) there exist $\lambda \in R^n$ and $\mu \in R^m$ such that $\lambda \geqslant 0$

$$\mu \geqq 0, \; \sum_{j=1}^{n} \lambda_j \nabla F_j(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) = 0,$$

and

$$\mu_i g_i(x^*) = 0; \; i = 1, 2, \ldots, m.$$

The following theorem generalizes Theorem 4.10 given in Chankong and Haimes[3] in the sense that the convexity conditions are weakened.

Theorem 2.5.1—Let x^* satisfy the KTCN for VOP ((1.1.1), (1.1.2)). Assume g and f are quasiconvex and for at least one k for which $\lambda_k > 0$ in the KTCN the corresponding f_k is strictly pseudoconvex at x^*. Then $x^* \in x^*$,

PROOF : Since x^* satisfies the KTCN for the VOP

$$\sum_{j=1}^{n} \lambda_j \nabla f_j(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) = 0.$$

Let $\lambda_k > 0$ and let the corresponding f_k be strictly pseudoconvex at x^*. Then the above equation can be rewritten as

$$\nabla f_k(x^*) + \sum_{\substack{j=1 \\ j\neq 1}}^{n} \lambda_j^0 \nabla f_j(x^*) + \sum_{i=1}^{m} \mu_i^0 \nabla g_i(x^*) = 0$$

where $\lambda_j^0 \geqq 0$ and $\mu_i^0 \geqq 0$ for $j = 1, 2, \ldots, n$; $j \neq k$ and $i = 1, 2, \ldots, m$. This together with other conditions above are equivalent to the Kuhn-Tucker conditions for optimality of P_k $(\epsilon^*_{(k)})$. The rest of the argument is similar to that in the proof of Theorem 2.3.2 (See Remarks in the paragraph immediately preceding theorem 2.3.3).

2.6. *Sufficient Conditions for Proper Noninferiority*

Geoffrion[4] reformulated the definition of proper noninferiority given by Kuhn-Tucker[8]. The following is the definition given by Geoffrion[4].

Definition —A noninferior solution of the VOP, x^*, is said to be a proper noninferior solution if there exists a scalar $M > 0$ such that for each i, $i = 1, 2, \ldots, n$ and each $x \in X$ satisfying $f_i(x) < f_i(x^*)$ there exists at least one $j \neq i$ with $f_j(x) > f_j(x^*)$ and

$$\frac{f_i(x) - f_i(x^*)}{f_j(x^*) - f_j(x)} \leqq M.$$

We have the following theorem on proper noninferiority which is a slight generalization of Theorem 4.13b of Chankong and Haimes[3].

Theorem 2.6.1—Let S be an open set in R^N. Let $x^* \in S$ solve $P_k(\epsilon^0_{(k)})$ and let $(x^*, \bar{v}, w)$ satisfy the Kuhn-Tucker conditions (2.1.12) and let the Kuhn-Tucker multipliers w_j associated with the constraints $f_j(x) \leqq f_j(x^0)$, $j \neq k$, be strictly positive. Let

(i) $w^T f$ be pseudoconvex (η-pseudoconvex) at x^* and g_J be quasiconvex (η-quasiconvex for the same function η) at x^*,

or

(ii) $w^T f$ be pseudoconvex (η-pseudoconvex) at x^* and $\bar{v}_J^T g_J$ be quasiconvex (η-quasiconvex for the same, function η) at x^*,

or

(iii) $w^T f + \bar{v}_J^T g_J$ be pseudoconvex or η-pseudoconvex at x^*. Then x^* is a proper noninferior solution.

Proof : From Lemma 2.1.5 it follows that x^* solves $P(w)$. We have the Kuhn-Tucker multipliers w_j to be strictly positive. The theorem now follows from well known result of Geoffrion[4] that if x^* solves $P(w)$ for some $w > 0$, then x^* is a proper noninferior solution of the VOP.

3. Duality Theory in Nonlinear Multiobjective Programming

Let Γ be the set of $m \times n$ real matrices. Consider the set of ordered pairs of vectors (η, ξ), $\eta \in R^n$, $\xi \in R^m$, defined by

$F = \{(\eta, \xi) \in R^{n+m} : \eta \geqq f(x), \xi \geqq g(x)$, for some $x \in R^N\}$, and the multiple criteria optimization problem (MCOP)

$$D(\pi) : EF(D(\pi)) = \min_{(\eta,\xi)\in F} \{\eta + \pi^T \xi\}$$

where $\pi \in \Gamma$, and $EF(D(\pi))$ denote the set of noninferior or efficient values of the objective function of Problem $D(\pi)$ Let $S = R^N$. Let

$$Y = \bigcup_{\pi\in\Gamma} EF(D(\pi)).$$

The dual problem to the VOP ((1 1.1), (1.1.2)) is defined to be MCOP

$$(D) : EF(D) = \underset{\rho \in Y}{\text{Max}} \ \rho$$

where $EF(D)$ denotes the set of noninferior or efficient values of the objective function of Problem (D).

We adopt the following conventions.

(a) For a minimization [maximization] MCOP we define EF (MCOP) $= (-\infty, \ldots, -\infty)$ $[(+\infty, \ldots; +\infty)]$ whenever for every $x^* \in X$ there exists, a sequence $\{x^j\}_{j=1}^{\infty}$, with $x^j \in X$ such that

$$f(x^{j+1}) \leqslant f(x^j) \leqq f(x^*) \quad [f(x^{j+1}) \geqslant f(x^j) \geqq f(x^*)]$$

and for some $k \in \{1, 2, \ldots, n\}$

$$\lim_{j \to \infty} f_k(x^j) = -\infty \, [+\infty].$$

(b) Whenever the feasible set of minimization [maximization] MCOP is empty, we define

$$EF\,(\text{MCOP}) = (+\infty, \ldots, +\infty) \, [(-\infty, \ldots, -\infty)].$$

A matrix π is said to be feasible in Problem (D) if $EF(D(\pi)) \neq (-\infty, \ \ , -\infty)$.

The following results are given by Bitran[2].

(3A) For any feasible matrix π in Problem (D) there is an $w \in R^n$ such that $w > 0$, $\pi w \geqq 0$.

(3B) Let x^* be feasible in the VOP ((1.1.1), (1.1.2)) $f(x^*) \in EF(D(\pi))$ for some $\pi \in \Gamma$. Then x^* is a noninferior solution of the VOP.

(3C) Assume that $x^* \in X = \{x \in R^N : g(x) \leqq 0\}$ is a noninferior solution of the following linear approximation multiple criteria optimization problem (LAP) to the VOP at x^*.

$$(\text{LAP}) : \text{Min} \{x^T \nabla f(x^*) : -(x - x^*)^T \nabla g_I(x^*) \geqq 0\}$$

where $I = \{i : g_i(x^*) = 0\} = \{1, 2, \ldots, q\}$ say; $I \subseteq \{1, 2, \ldots, m\}$.

The index set I corresponds to the constraints active at x^*. Then there exist $q \times n$ matrices U^1, U^2, U^*, with

$$U^* = U^1 + U^2$$

such that

$$\lambda^T [\nabla f(x^*) + \nabla g_I(x^*) U^*] - t'^T u^* \leqslant 0$$

$$\lambda \in R^N, \ t' \in R^q, \ t' \geqq 0 \qquad \ldots(3.1.1)$$

has no solution and

$$-x^{*T} \nabla f(x^*) = x^{*T} \nabla g_I(x^*) U^1 + e_1^{0^T} U^2 \qquad \ldots(3.1.2)$$

where $e_1^0 \in R^q$ and $e_1^{0^T} = (1, 0, \ldots, 0)$.

The above result 3.C has been given by Bitran[2] as a consequence of a similar result obtained by Isermann[6] for MOLPP.

It may be noted that construction of (LAP) at x^* such that x^* itself is its noninferior solution may pose a separate problem.

Bitran[2] has also extended the result so as to include, the constraints nonactive at x^*,

Let $I' = \{i : g_I(x^*) < 0\}$ be the index set corresponding to the constraints nonactive at x^*.

Define

$$\pi^i = \binom{U^i}{0}; \; i = 1, 2, * \qquad \ldots(3.1.3)$$

where the zero matrix in π^i is of order $(m-q) \times n$ and corresponds to the constraints nonactive at x^*. Clearly π^1, π^2, π^* are $m \times n$ matrices. From the relation between U^*, U^1 and U^2, it follows that $\pi^* = \pi^1 + \pi^2$.

We can now rewrite (3.1.1) and (3.1.2), respectively, as

$$\lambda^T [\nabla f(x^*) + \nabla g(x^*) \pi^*] + t^T \pi^* \leqslant 0$$

$$\lambda \in R^N, t \in R^m, t \geqq 0 \qquad \ldots(3.1.4)$$

has no solution and

$$-x^{*^T} \nabla f(x^*) = x^{*^T} \nabla g(x^*) \pi^1 + e_1^I \pi^2 \qquad \ldots(3.1.5)$$

$$e_1 \in R^m, e_1^T = 1, 0, \ldots, 0.$$

The above result can be stated in a more general setting as follows :

Let $x^* \in X = \{x \in R^N : g(x) \leqq 0\}$ be a noninferior solution to the following LAP associated with the VOP.

$$(\text{LAP})^* : \text{Min} \{x^T \nabla f(x) : -(x - x^*)^T \nabla g(x^*) \geqq 0\}.$$

Then there exist $m \times n$ matrices π^1, π^2 and π^* such that $\pi^* = \pi^1 + \pi^2$ and such that

$$\lambda^T [\nabla f(x^*) + \nabla g(x^*) \pi^*] + t^T \pi^* \leqslant 0,$$

$$\lambda \in R^N, t \in R^m, t \geqq 0 \qquad \ldots(3.1.6)$$

has no solution and

$$-x^{*^T} \nabla f(x^*) = x^{*^T} \nabla g(x^*)\,\pi^1 + \alpha^T \pi^2 \qquad \text{...(3.1.7)}$$

for some $\alpha \in R^m$.

We now state the following :

Weak Duality Theorem[2]—For any matrix π feasible in Problem (D) and any x feasible in the VOP, we have

$$f(x) \nleq z \text{ for all } z \in EF(D(\pi)).$$

Combining the result (3A), (3B) and the Weak Duality Theorem above, Bitran[2] has given the following important result.

(3D) Let x be feasible in the VOP, $f(x) \in EF(D(\pi))$ for some $\pi \in \Gamma$. Then $f(x) \in EF(D)$, x is a noninferier solution of the VOP and π is a noninferior solution of Problem (D).

3.1. *Relationships Among the Solution of the* (LAP), *Noninferior Solution of the* VOP, *and Noninferior Solution of Problem* (D)

We now prove the following theorem with the notation and results developed above.

Theorem 3.1.1—If a noninferior solution x^* of the linear approximation (LAP) MCOP to the VOP at x^* is feasible for the VOP, f is pseudoconvex and g_I is differentiable and quasiconvex at x^*, then $x^* \in X^*$.

PROOF : Assume $x^* \notin X^*$. Then there exists an $\hat{x} \in X$, $\hat{x} \neq x^*$ such that $f(\hat{x}) \leqslant f(x^*)$. Since f is pseudoconvex at x^*, this implies

$$(\hat{x} - x^*)^T \nabla f(x^*) \leqslant 0.$$

Since $x^* \in X$, $g_I(\hat{x}) \leqq g_I(x^*) = 0$, and further since g_I is quasiconvex at x^*,

$$(\hat{x} - x^*)^T \nabla g_I(x^*) \leqq 0.$$

So

$$\hat{x} \in \{x \in R^N : -(x - x^*)^T \nabla g_I(x^*) \geqq 0 \text{ and}$$
$$(x - x^*)^T \nabla f(x^*) \leqslant 0\}.$$

This contradicts the statement that x^* is a noninferior solution of (LAP). Consequently every point $\xi \in X$, which is noninferior solution of (LAP) at ξ itself, is a noninferior solution of the VOP.

Remark : The above theorem is an improvisation of earlier known result [Bitran[2], Proposition 2.6 (i)], where the VOP is a maximization problem and f and g

are assumed to be concave and differentiable on R^N. Here we have weakened the convexity/concavity assumptions.

We now give the converse of the above theorem.

Theorem 3.1.2—Let x^* be a noninferior solution of the VOP. Let f be pseudoconcave and g be pseudoconcave at x^*. Then x^* is a noninferior solution of the (LAP)*:

$$\text{(LAP)}^* : \text{Min}\,\{x^T \nabla f(x^*) : -(x - x^*)^T \nabla g(x^*) \geqq 0\}.$$

PROOF : Assume that x^* is not a noninferior solution of (LAP)*; then there exists an $\hat{x} \neq x^*$ such that

$$\hat{x}^T \nabla f(x^*) \leqslant x^{*^T} \nabla f\ (x^*)$$

i. e. $$(\hat{x} - x^*)^T \nabla f_i\,(x^*) \leqslant 0 \qquad \ldots(3.1.8)$$

subject to $(\hat{x} - x^*)^T \nabla g\,(x^*) \leqq 0.$

Since g is pseudoconcave at x^*, this implies $g\,(\hat{x}) \leqq g\,(x^*) \leqq 0$, i. e. $\hat{x} \in X$.

Since f is pseudoconcave and hence quasiconcave, for all components i of the L.H.S. of the vector inequality (3.1.8) for which $(\hat{x} - x^*)^T \nabla f_i\,(x^*) < 0$ holds, we shall have $f_i\ (\hat{x}) < f_i\ (x^*)$. On the other hand for all other components j we shall have from the same vector inequality $(\hat{x} - x^*)^T \nabla f_j\,(x^*) = 0$. Since f is pseudoconcave, $f_j\,(\hat{x}) \leqq f_j\,(x^*)$.

Combining these we shall have

$$f\,(\hat{x}) \leqslant f\,(x^*). \qquad \ldots(3.1.9)$$

But (3.1.9) contradicts that x^* is a noninferior solution of the VOP. Hence x^* is a noninferior solution of the (LAP)*.

Let x^* be a noninferior solution of (LAP)*. Then from result (3C) it follows

$$\lambda^T\,[\nabla f\,(x^*) + \nabla g\,(x^*)\,\pi^*] + t^T\,\pi^* \leqslant 0$$
$$\lambda \in R^N,\ \ t \in R^m,\ t \geqq 0 \qquad \ldots(3.1.6)$$

has no solution and

$$-\,x^{*T}\,\nabla f\,(x^*) = x^{*^T} \nabla g\,(x^*)\,\pi^1 + \alpha^T\,\pi^2, \text{ for some } \alpha \in R^m. \qquad \ldots(3.1.7)$$

By applying Motzkin's theorem of the alternative[12] to (3.1.6) it follows that there exists a $w_0 \in R^n$, $w_0 > 0$ such that

$$[\nabla f(x^*) + \nabla g(x^*)\pi^*] w_0 = 0 \quad \text{...(3.1.10)}$$

$$\pi^* w_0 \geqq 0. \quad \text{...(3.1.11)}$$

First, we note from the definition of π^* in (3.1.3), that

$$\pi^*_{ji} g_i(x^*) = 0, j = 1, 2, \ldots, m;$$

$$i = 1, 2, \ldots, n. \quad \text{...(3.1.11a)}$$

We shall now make use of w_0 determined in (3.1.10) in the following theorem.

Theorem 3 1.3—Assume that x^* is a noninferior solution of (LAP)* such that $x^* \in X$. Assume that $w_0^T [f + \pi^{*T} g]$ is η-pseudoconvex at x^* where π^* is defined in (3.1.3). Then (i) $f(x^*) \in EF(D(\pi^*))$ (ii) $f(x^*) \in EF(D)$ (iii) π^* is a noninferior solution of Problem (D).

PROOF : Note that, by definition,

$$D(\pi^*) = \min_{(\eta, \xi) \in F} \{\eta + \pi^{*T}\xi\}.$$

Assume that $f(x^*) \notin EF(D(\pi^*))$. Then there exists an $\bar{x} \in R^N$, $(\bar{\eta}, \bar{\xi}) \in R^{n+m}$ such that

$$\bar{\eta} \geqq f(\bar{x}), \bar{\xi} \geqq g(\bar{x}) \text{ and}$$

$$\bar{\eta} + \pi^{*T}\bar{\xi} \leqslant f(x^*) = f(x^*) + \pi^{*T} g(x^*) \quad \text{...(3.1.12)}$$

since $\pi^{*T} g(x^*) = 0$.

Multiplying (3.1.12) on the left by w_0^T, and noting $w_0 > 0$, we get

$$w_0^T \bar{\eta} + w_0^T \pi^{*T}\bar{\xi} < w_0^T f(x^*) + w_0^T \pi^{*T} g(x^*).$$

Also, from $\bar{\eta} \geqq f(\bar{x})$ and $\bar{\xi} \geq g(\bar{x})$, we get

$$w_0^T f(\bar{x}) + w_0^T \pi^{*T} g(\bar{x}) \leqq w_0^T \bar{\eta} + w_6^T \pi^{*T}\bar{\xi}.$$

Hence

$$w_0^T f(\bar{x}) + w_0^T \pi^{*T} g(\bar{x}) < w_0^T f(x^*) + w_0^T \pi^{*T} g(x^*). \quad \text{...(3.1.13)}$$

Since $w_0^T\ [f + \pi^{*T} g]$ is η-pseudoconvex at x^* from (3.1.10) we get

$$w_0^T f(\bar{x}) + w_0^T\ \pi^{*T} g(\bar{x}) \geqq w_0^T f(x^*) + w_0^T\ \pi^{*T} g(x^*).$$

This contradicts (3.1.13).

Therefore $f(x^*) = f(x^*) + \pi^{*T} g(x^*) \in EF(D(\pi^*))$. This proves (i), (ii), (iii) follows from result 3. D.

Remark 3.1.1 : Bitran[2] (Proposition 2.6 (ii)) considered a similar problem of maximization where f and g were assumed to be concave and differentiable on R^N. In the Theorem 3.1.3 above it has been found enough to assume $w_0^T\ [f + \pi^{*T} g]$ to be η-pseudoconvex at x^*. In this sense Proposition 2.6 (ii) of Bitran[2] is a particular case of Theorem 3.1.3 above.

Remark 3.1.2 : Theorems similar to Theorems 3.1.1—3.1.3 can be formulated in a straight forward manner so as to apply to situations where not all components of the constraint function g are differentiable.

3.2. *Direct Duality*

We begin the discussion of direct duality theory with the following theorem.

Theorem 3.2.1 Let x^* be a noninferior solution of the VOP. Assume that (x^*, v, w), $v \in R^m$, $w \in R^n$ satisfies

$$\nabla f(x^*)\ w + \nabla\ g(x^*)\ v = 0 \qquad \text{...(3.2.1a)}$$

$$v^T g\ (x^*) = 0 \qquad \text{...(3.2.1b)}$$

$$v \geqq 0,\ w \geqslant 0,\ \sum_{i=1}^{n} w_i = 1. \qquad \text{...(3.2.1c)}$$

Assume that $w^T f + v^T g$ is η-strictly yseudoconvex at x^*. Then there exists a $\pi^0 \in \Gamma$ such that $f(x^*) \in EF(D(\pi^0))$ and $\pi^0 \in EF(D))$.

PROOF : In the vector w some of the component may be zero. Define π^0 as the $m \times n$ matrix whose ith column is equal to zero or v according as $w_i = 0$ or $w_i \neq 0$. Then $\pi^0 w = v$. Assume that $f(x^*) \notin EF(D(\pi^0))$. Then there exists $(\eta, \xi) \in F$ such that

$$\eta + \pi^{0^T}\ \xi \leqslant f(x^*).$$

Hence

$$w^T\eta +\ \pi^T\ \pi^{0^T}\ \xi = w^T\ \eta + v^T\ \xi \leqslant\ w^T f(x^*)$$

The rest of the proof of this theorem is similar to that of Theorem 3.1.3.

We now use the concept of selective duality and obtain the following direct duality theorem. Assume that a nonempty subset Δ of $\{1, 2, ..., n\}$ is such that for each $j \in \Delta$, f_j is differentiable at x^*. Similarly, suppose a nonempty subset Δ' of $\{1, 2, ..., m\}$ is such that for each $i \in \Delta'$, g_i is differentiable at x^*. Let f_Δ, $g_{\Delta'}$ be the subvectors of f and g corresponding to the index sets Δ and Δ' respectively.

Theorem 3.2.2—Let x^* be a noninferior solution of the VOP. Assume that $(x^*, v_\Delta, w_\Delta)$ satisfies

$$\nabla f_{\Delta}(x^*)\, w_{\Delta} + \nabla g_{\Delta'}(x^*)\, v_{\Delta'} = 0 \qquad ...(3.2.2a)$$

$$v^T_{\Delta'}\, g_{\Delta'}(x^*) = 0 \qquad ...(3.2.2b)$$

$$v_{\Delta'} \geqq 0,\ w_\Delta \geqslant 0,\ \sum_{i \in \Delta} (w_{\Delta})_i = 1. \qquad ...(3.2.2c)$$

Assume that $w^T_{\Delta} f_{\Delta} + v^T_{\Delta'} g_{\Delta'}$ is η-strictly pseudoconvex at x^*. Then there exists a $\pi^0 \in \Gamma$ such that $f(x^*) \in EF(D(\pi^0))$ and $\pi^0 \in EF(D)$.

Proof : Define $v \in R^m$, $v = \begin{bmatrix} v_\Delta \\ 0 \end{bmatrix}$ where the $(m - |\Delta'|)$- zero vector corresponds to the constraint functions nondifferentiable at x^*. Also define

$w = \begin{bmatrix} w_\Delta \\ 0 \end{bmatrix} \in R^n$ where the $(n - |\Delta|)$-zero vector corresponds to the objective functions nondifferentiable at x^*.

Define π^0 as the $m \times n$ matrix with the ith column equal to zero or v according as $w_i = 0$ or $w_i \neq 0$.

Assume $f(x^*) \notin EF(D(\pi^0))$. Then there exists $(\eta, \xi) \in F$ such that

$$\eta + \pi^{0^T} \xi \leqslant f(x^*)$$

and $\eta \geqq f(\bar{x}), \xi \geqq g(\bar{x})$ for some $\bar{x} \in R^N$.

Hence

$$w^T \eta + w^T \pi^{0^T} \xi \leqq w^T f(x^*),$$

which implies

$$w^T \eta + v^T \xi \leqq w^T f(x^*).$$

This further gives

$$w^T f(\bar{x}) + v^T g(\bar{x}) \leqq w^T \eta + v^T\xi \leqq w^T f(x^*)$$

from which we get

$$w_{\Delta}^{T} f_{\Delta}(\bar{x}) + v_{\Delta'}^{T} g_{\Delta'}(\bar{x}) \leqq w_{\Delta}^{T} f_{\Delta}(x^*) + v_{\Delta'}^{T} g_{\Delta'}(x^*)$$

and as $w_{\Delta}^{T} f_{\Delta} + v_{\Delta'}^{T} g_{\Delta'}$ is assumed to be η-strictly pseudoconvex at x^* this implies

$$\eta^T(\bar{x}, x^*)[\nabla f_{\Delta}(x^*) w_{\Delta} + \nabla g_{\Delta'}(x^*) v_{\Delta'}] < 0.$$

But (3.2.2a) implies

$$\eta^T(\bar{x}, x^*)[\nabla f_{\Delta}(x^*) w_{\Delta} + \nabla g_{\Delta'}(x^*) v_{\Delta'}] = 0$$

which is a contradiction. Hence $f(x^*) \in EF(D(\pi^0))$ and also $f(x^*) \in EF(D)$ follows from result 3D.

Remark 3.2.2 : As a special case of the above theorem we obtian the usual direct duality theorem.

3.3. *Saddle Point Duality, Stability and Kuhn-Tucker Conditions.*

Definition 3.3.1—A pair (x^*, π^*) is said to be a saddle point of the vector valued Lagrangian

$$L(\eta, \xi, \pi) = \eta + \pi^T \xi$$

if

$$f(x^*) + \pi^T g(x^*) \not\geq f(x^*) + \pi^{*T} g(x^*) \not\geq \eta + \pi^{*T} \xi \qquad \ldots(3.3.1)$$

for all $(\eta, \xi) \in F$ and all π such that

$$\pi^T \xi \leqslant 0, \ \xi \geqq 0,$$

has no solution.

Definition 3.3.2—The perturbation point-to-set map $v : R^m \to \mathscr{P}(R^n)$ is defined as

$$v(y) = EF(P_y)$$

where $EF(P_y)$ is defined through the problem (P_y) given by

$$(P_y) : EF(P_y) = \min \{f(x) : x \in X_y\}$$

$$X_y = \{x \in R^N : g(x) \leqq y\}.$$

That is, $v(y)$ is the image of $EF(X_y)$ under $f(\,.\,)$. Here the constraints $g(x) \leqq 0$ are perturbed to $g(x) \leqq y$, the vector y being the perturbation. Clearly, therefore, $X = X_0$.

Definition 3.3.3—The VOP is said to be M-stable at $x^* \in X$ if there is a $n \times m$ matrix M such that

$$v(y) \cap \{(f(x^*) - My) + R_{-}^{n}\} = \phi \text{ for all } y \in R^m \qquad \ldots(3.3.2)$$

where

$$R^n_- = \{x \in R^n : x \leqslant 0\}.$$

Definition 3.3.4—Kuhn-Tucker conditions :

The Kuhn-Tucker conditions for a pair (x^*, π^*) with $x^* \in R^N$ and $\pi^* \in \Gamma$ can be stated as follows :

$$\left.\begin{array}{l} \lambda^T [\nabla f(x^*) + \nabla g(x^*) \pi^*] + t^T \pi^* \leqslant 0 \\ \lambda \in R^N, t \in R^m, t \geqq 0 \end{array}\right\} \text{has no solution} \quad \ldots (3.3.3a)$$

$$\pi^{*T} g(x^*) = 0; \quad \ldots(3.3.3b)$$

$$g(x^*) \leqq 0. \quad \ldots(3.3.3c)$$

The Kuhn-Tucker conditions (3.3.3) can also be stated as follows :

$$\begin{array}{l} [\nabla f(x^*) + \nabla g(x^*) \pi^*] w_0 = 0, \ w_0 \geqslant 0, \\ \pi^* w_0 \geqq 0 \\ \pi^{*T} g(x^*) = 0 \\ g(x^*) \leqq 0. \end{array} \quad \ldots (3.3.4)$$

The next theorem gives the relation among the Kuhn-Tucker conditions, a noninferior solution of the VOP and a noninferior solution of Problem (D). We state it without proof since it is a paraphrase of Theorem 3.1.3 and result (3D). This theorem is an improvisation of Proposition 3.6 of Bitran[2] in the sense that convexity assumptions are weakened.

Theorem 3.3.1—Assume that the Kuhn-Tucker conditions (3.3.4) hold at (x^*, π^*, w_0) and $w_0^T [f + \pi^{*T} g]$ is η-pseudoconvex at x^*. Then $f(x^*) \in EF(D(\pi^*))$, x^* is a noninferior solution of the VOP, $f(x^*) \in EF(D)$ and x^* is a noninferior solution of Problem (D).

The Kuhn-Tucker conditions for the case when objective functions and constraint functions are not all differentiable can be formulated as follows. Let $\Delta, \Delta', f_\Delta, g_\Delta$, have the same meaning as defined earlier.

Write

$$\pi^* = \begin{array}{c c} & \begin{array}{cc} |\Delta| & n - |\Delta| \end{array} \\ \begin{array}{c} |\Delta'| \\ m - |\Delta'| \end{array} & \begin{bmatrix} \pi^*_{\Delta'\Delta} & 0 \\ 0 & 0 \end{bmatrix} \end{array}. \quad \ldots(3.3.5)$$

The Kuhn-Tucker conditions corresponding to (3.3.4) can now be stated as follows.

$$[\nabla f_{\Delta}(x^*) + \nabla g_{\Delta'}(x^*)\ \pi^*_{\Delta'\Delta}]\ w_{\Delta} = 0,\ w_{\Delta} > 0,$$

$$\pi^*_{\Delta'\Delta}\ \ w_{\Delta} \geqq 0,$$

$$\pi^*_{\Delta\Delta}\ g_{\Delta'}(x^*) = 0 \qquad \ldots(3.3.6)$$

$$g_{\Delta'}(x^*) \leqq 0.$$

We now state a theorem involving LAP's associated with the VOP where not all components of f and g are differentiable.

Theorem 3.3.2—If $x^* \in X$ solves the following linear approximation problem (LAP_{Δ}) to the VOP at x^* :

$$(LAP_{\Delta}) : \text{Min } \{x^T \nabla f_{\Delta}(x^*) : -(x - x^*)^T \nabla g_{I_{(\Delta')}}(x^*) \geqq 0$$

then there is a $|\Delta'| \times |\Delta|$ matrix $\pi^*_{\Delta'\Delta}$ and w_{Δ} such that $(x^*, \pi^*_{\Delta'\Delta}, w_{\Delta})$ satisfies the Kuhn-Tucker conditions (3.3.6).

The following theorem is straight forward extension of Proposition 3.5 of Bitran[2].

Theorem 3.3.4—Assume that the Kuhn-Tucker conditions (3.3.4) hold at a pair (x^*, π^*) and $w_0^T\ [f + \pi^{*T} g]$ is η-pseudoconvex at x^*, then (i) the VOP is M-stable at x^* with $M(x^*) = \pi^*$ and (ii) (x^*, π^*) is a saddle point of $L(\eta, \xi, \pi)$

Theorem 3.3.5—Let (i) (x^*, π^*) be a saddle point of $L(\eta, \xi, \pi)$ or (ii) the VOP be M-stable at x^* with $M(x^*) = \pi^*$ and $\pi^{*T} g(x^*) = 0$. Let f be pseudoconcave and let the constraint g bo pseudoconcave at x^*. Then the Kuhn-Tucker condition (3.3.5) hold at the pair (x^*, π^*).

PROOF : From Proposition 3.5 of Bitran[2] we have x^* is a noninferior solution of the VOP. Then from Theorem 3.1.2 we have x^* is a noninferior solution of $(LAP)^*$ of Theorem 3.1.2, and consequently (x^*, π^*) satisfies the Kuhn-Tucker conditions (3.3.4).

The following theorem gives the conditions under which a solution of the weighting problem $P(w)$ satisfies the Kuhn-Tucker conditions (3.3.4).

Theorem 3.3.6—Let x^* solve $P(w_0)$ where $w_0 > 0$. Assume w_0^T is quasiconcave and differentiable at x^* and g_I is quasiconvex and differentiable at x^*. Then x^* satisfies the Kuhn-Tucker conditions (3.3.4).

PROOF : Since $w_0^T f$ is quasiconcave and differentiable at x^* and g_I is quasiconvex and differentiable at x^*, we get x^* solves the linear approximation to problem $P(w_0)$ at x^*, i. e.,

$$\text{Min } \{x^T \nabla f(x^*) \, w_0 : -(x - x^*)^T \nabla g_i(x^*) \geqq 0\}.$$

Suppose x^* does not solve the linear approximation to the VOP at x^* i. e. there exists a $\hat{x} \in X$, $\hat{x} \neq x^*$ such that

$$\hat{x}^T \nabla f(x^*) \leqslant x^{*T} \nabla f(x^*)$$

or
$$(\hat{x} - x^*)^T \nabla f(x^*) \leqslant 0.$$

As $w_0 > 0$ this implies $(\hat{x} - x^*)^T \nabla f(x^*) \, w_0 < 0$.

This contradicts that x^* solves the linear approximation to Problem $P(w_0)$ at x^*. Hence x^* solves the linear approximation to the VOP at x^*. This implies there is $\pi \in \Gamma$ such that (x^*, π^*) satisfies the Kuhn-Tucker conditions (3.3.4).

CONCLUDING REMARKS

In this paper an attempt has been made to investigate the extent to which convexity/concavity assumptions on the objective functions or constraint functions in a VOP can be relaxed to weaker assumptions of pseudoconvexity/pseudoconcavity, quasiconvexity/quasiconcavity, etc. The problems of necessary and sufficient optimality conditions, some aspects of duality theory such as direct duality and selective duality, have been analyzed with this idea in mind. It seems the converse duality theory is not amenable to such processes of weakening the underlined convexity/concavity assumption.

REFERENCES

1. M. S. Bazaraa and C. M. Shetty, *Nonlinear Programming Theory and Algorithms*. John Wiley and Sons, 1979.
2. G. R. Bitran, *JOTA*, 35 (1981), 367-401.
3. V. Chankong and Y. Y. Haimes, *Multi-objective Decision Making Theory and Methodology*. North Holland Publ. Co., New York, 1983.
4. A. M. Geoffrion, *J. Math. Anal. Appl.* 22 (1968), 618-630.
5. M. A. Hanson, *J. Math. Anal Appl.* 80 (1981), 545-650.
6. H. Isermann, *The Relevance of Duality in Multiple Objective Linear Programming*, *TIMS Studies in the Management Sciences*, Vol. 6, North Holland Publishing Company, New York, 1977.
7. R. N. Kual and S. Kaur, *J. Math. Anal. Appl.* 105 (1985), 104-112.
8. H. W. Kuhn and A. W. Tucker, *Nonlinear Programming in Proceedings of the Second Berkely Symposium of Maths., Stats. and Probability*. University of California, Berkeley, 1951, 481-492.
9. J. G. Lin, *Many Person Optimization and Decentralized Control*. Plenum, New York, 1976, 117-138.

10. J. G. Lin, *JOTA* 21 (1977), 505-521.
11. D. G. Mahajan, Doctoral Thesis I. I. T. Powai, Bombay, 1977.
12. O. L. Mangasarian, *Nonlinear Programming*. McGraw-Hill Book Co., Inc. New York, 1969.
13. G. P. Mc.Cormick, *SIAM J. Appl. Math.* 15 (1967), 641-652.
14. J. Ponstein, *SIAM Rev.* 9 (1967), 115-119.

Indian J. pure appl. Math., 20 (1) : 40-48, January, 1988

ON AUTOMORPHISMS OF FREE GROUPS

A. G. A. E Abdel-Gawad

Department of Mathematics, Faculty of Science, Suez Canal University, Ismaila Egypt

(*Received* 6 *August* 1987; *after revision* 4 *January* 1988)

In this paper, we study free groups and the automorphism of free groups in connection with a graph. Basic definitions and some results are given. The relevant main theorem "A graph satisfying conditions G1-3 of Gersten gives a based graph satisfying (2′)" is proved.

Introduction

In sections one, and two the basic definitions and some results on graphs and free groups are given. Also, we give the relation between an automorphism of a free group and a graph. In section three, the conditions of based graphs and the definition of the minimal based graph are given.

The main future of this work is to prove that a based graph satisfying conditions G1-3 of Gersten[2] gives a based graph satisfying (2′) which we consider it to be a generalization of Abdel-Gawad[1].

1. Graphs

1.1. A graph Γ consists of two disjoint sets, a non-empty set of vertices $V(\Gamma)$ and a set of edges $E(\Gamma)$, together with a function $t : E(\Gamma) \to E(\Gamma)$ denoted by $e \to \bar{e} \neq e$ (so $\bar{\bar{e}} = e$). Let $i(e) = t(\bar{e})$, where $i(e)$ is called the initial point of e and $t(\bar{e})$ the terminal point of $\bar{e}$ ($\bar{e}$ is called the inverse of e).

If e is in $E(\Gamma)$ and $i(e) = u$; $t(e) = v$, we say e joins u and v and write $e : u \to v$.

1.2. Apath p in the graph Γ is an ordered n-tuple $(n \geqslant 1)$ $p = (e_1, \ldots, e_n)$; $e_i \in E(\Gamma)$; $i = 1, \ldots, n$; such that $t(e_j) = i(e_{j+1})$; $1 \leqslant j \leqslant n-1$; $i(e_1)$ and $t(e_n)$ are called the initial and terminal points of p, written $i(p)$ and $t(p)$ respectively ($\bar{p} = (\bar{e}_n, \ldots, \bar{e}_1)$).

If two paths $p = (e_1, \ldots, e_n)$ and $p' = (e'_1, \ldots, e'_m)$ are such that $i(p') = t(p)$, then their composite $p.p'$ is defined by $p.p' = (e_1, \ldots, e_n, e'_1, \ldots, e'_m)$.

1.3. An elementary reduction $p \searrow p'$ of paths is a reduction by which $p = (e_1, \ldots, e_n)$ is replaced by

$p' = (e_1, \ldots, e_{j-1}, e_{j+2}, \ldots, e_n)$ if $e_{j+1} = \bar{e}_j$.

1.4. The path p is called reduced if it admits no elementary reductions.

1.5. Two paths p and p' are called equivalent, written $p \sim p'$ if there is a finite sequence of paths $p = p_1, \ldots, p_m = p'$ such that either $p_j \curvearrowright p_{j+1}$ or $p_{j+1} \curvearrowright p_j$ for $1 \leqslant j \leqslant m-1$.

Proposition 1.1— (a) Each path is equivalent to a unique reduced path; (b) the operation of composition of paths is consistent with equivalence[2,3,9,10].

1.6. If p_1 and p_2 are reduced paths, then $p_1 p_2$ denotes the reduced path equivalent to p_1, p_2, if defined.

1.7 A morphism[2,3] of graphs $f: \Gamma \to \Gamma'$ is a map f taking each vertex to aertex, and each edge to an edge or a vertex such that $f(\bar{e}) = \overline{f(e)}$ and $t(f(e)) = f(t(e))$ where $\bar{v} = t(v) = v$.

1.8. A graph is called connected[9] if every pair of its vertices is joined by some path. In general, a graph is the disjoint union of its connected components.

1.9. A loop at a vertex v is a path p not necessarily reduced, with $i(p) = t(p) = v$.

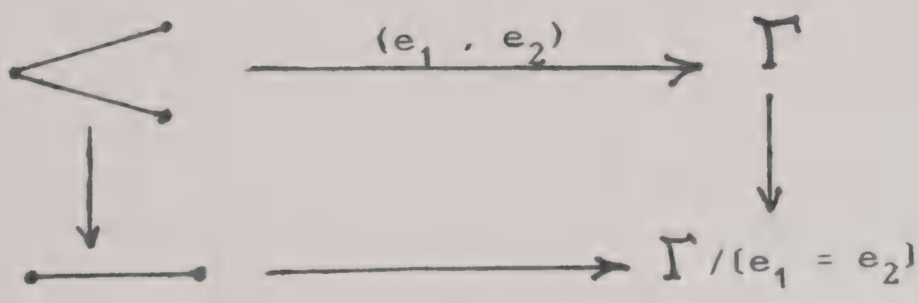
1.10. If the loop at a vertex v is reduced path, it will be called circuit.

1.11. A tree Serre[8] is a connected non-empty graph without circuits.

1.12. Let Γ' be a subgraph of a connected graph Γ. W say that Γ' is spanning if every pair of vertices of Γ is joined by at least one path Γ'.

1.13. A subgraph Γ' is a spanning tree or maximal subtree of Γ if Γ' is a tree and spanning.

1.14. A pair of edges (e_1, e_2) of Γ is said to be admissible Stallings[9] if $i(e_1) = i(e_2)$ and $e_1 \neq \bar{e}_2$. In this case we can identify $t(e_1)$ to $t(e_2)$, e_1 to e_2, $\bar{e}_1$ to $\bar{e}_2$ to obtain a graph denoted by $\Gamma/[e_1 = e_2]$. The morphism $\Gamma \to \Gamma/[e_1 = e_2]$ is called an edge fold.



1.15. Let Γ, Γ' be two graphs, an edge collapse is a morphism $f: \Gamma \to \Gamma'$ whose effect is to identify e, $\bar{e}$, $i(e)$ and $t(e)$ to one vertex.

2. Free Groups and Graphs Representing Automorphisms of Free Groups

2.1. *Free Groups*

Let $X = \{x_1^{\pm 1}, \ldots, x_n^{\pm 1}\}$ be a finite set whose elements are called letters. We define an equivalence ($\sim$) on words in X generated by $w_1 x_i^{\epsilon} x_i^{-\epsilon} w_2 \sim w_1 w_2$; $i = 1, \ldots, n$; $\epsilon = \pm 1$, where w_1, w_2 are arbitrary words in X, and $(x^{-1})^{-1} = x^{-\epsilon}$, where x_i is called the inverse of x_i^{ϵ}; $\epsilon = \pm 1$. If $w = x_{i_1}^{\epsilon_{i_m}}, \ldots, x_{i_m}^{\epsilon_{i_m}}$ then w^{-1} is defined to be $x_{i_m}^{-\epsilon_{i_m}} \ldots x_{i_1}^{-\epsilon_{i_1}}$.

The operation of juxtaposition is compatible with equivalence. Therefore if the equivalence class of w is denoted by $[w]$ then the operation on the classes is well defined.

$$[w].\,[w'] = [w.\,w'] \qquad \ldots(1)$$

where the R.H.S. product is juxtaposition.

$[w].\,[w^{-1}] = [w^{-1}].\,[w] = [w.w^{-1}] = [w^{-1}.w] = [e]$, where e is the empty word.

Definition 2.1.1—The free group F_X on X is formed by taking as elements all equivalence classes of words in X and as group product, the operation defined by (1).

Definition 2.1.2—A word in X is said to be reduced or freely reduced if it contains no consecutive pairs of the form $x_i^{\epsilon}\ x_i^{-\epsilon}$; $\epsilon = \pm 1$; $i = 1, \ \ , n$.

Lemma 2.1.1—Any equivalence class contains one and only one, reduced word[4].

Henceforth the elements of F_X will be considered as reduced words.

2.2. *Automorphisms of Free Groups*

Let X and F_X be defined. We denote by Aut F_X, the automorphism group of the free group F_X.

Automorphisms of type (1) Lyndon[7] are those which merely premute the letters in X subject to the restriction that if $x_i \to x_j$, then $x_i^{-1} \to x_j^{-1}$; $i \neq j$. Automorphisms[5] of type (2) are denoted uniquely by the symbols (A, a), where A is a proper subset of X and a, which is termed the pivot of the transformation, satisfies $a \in A$, $a^{-1} \notin A$.

(A, a) is defined on F_X by mapping each letter as follows $(A, a)\, a = a^{-1}$ and $(A, a)\ a^{-1} = a$, and for $x \in X$, $x \neq a$, a^{-1} we have

$$(A, a)x = \begin{cases} a^{-1} x a & ; \ x, x^{-1} \in A \\ xa & ; \ x \notin A,\ x^{-1} \in A \\ a^{-1} x & ; \ x \in A,\ x^{-1} \notin A \\ x & ; \ x, x^{-1} \notin A. \end{cases}$$

For $\alpha \notin X$, define $(A, \alpha) : F_X \to F_{X_{\cup\alpha}} \pm 1$ as a monomorphism given by

$$(A,\ \alpha)\ x\ =\ \begin{cases} \alpha^{-1}\ x\alpha & ;\ x, x^{-1} \in A \\ x\alpha & ;\ x \notin A, x^{-1} \in A \\ \alpha^{-1}x & ;\ \alpha \in A, x^{-1} \notin A \\ x & ;\ x, x^{-1} \notin A. \end{cases}$$

This is a paraphase of Abdel-Gawad[1].

Definition 2.2.1—A labelling is a map $f : E(\Gamma) \rightarrow F_Y$, such that $f(\bar{e}) = f(e)^{-1}$ for all $e \in E(\Gamma)$. If $p = e_1 \ldots e_n$ is a path in Γ then $f(p) = f(e_1) \ldots Xf(e_n)$.

Definition 2.2.2—The based graph Γ which represets (A, α) has two vertices A and A' where A' (the complement of A in X) is the base point, and a directed edge labelled $x^{\pm 1}$, for each $x^{\pm 1} \in X$ and for α, α^{-1} such that $i(x) \in A \Rightarrow x \in A$, $i(\alpha) \in A$, $t(\alpha) \in A'$. Call this labelling f; as example, let $A = \{x_1, x_2^{\pm 1}, x_3^{-1}\} \subset X = \{x_1^{\pm 1} \ldots, x_4^{\pm}\}$.

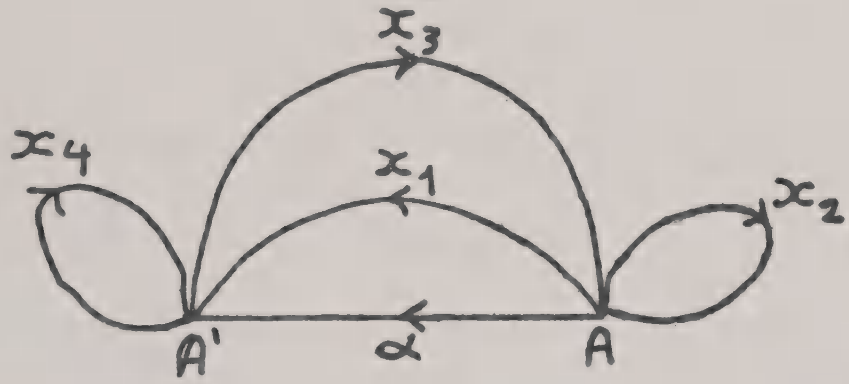


FIG. 1.

We now show how to construct from this graph Γ a graph representing (A, a). Define two labellings on Γ as follows. The first label f_1 of the edge labelled α is 1 and the other edges are the same as f. The second label f_2 is 1 on the directed edge with first label a, is a on the directed edge with first label 1, i.e. on edge labelled α in Γ, and is the same as the first label on the other edges.

Henceforth, we use (x, y) to denote on edge with first label x and second label y.

Proposition—For each $w \in Fx$ there exists a unique circuit p at A' such that $f_1(p) = w$. Moreover $f(p) = (A, \alpha)\, w$ and $f_2(p) = (A, a)\, w$.

PROOF : The result is clear for $w = x^{\pm 1}$ and is proved for arbitrary w by induction.

From the last example, if $a = x_1$; then

$(A, a) = (\alpha = 1).\ (x_1 \Leftrightarrow \alpha)\ (A, a)$ and $\alpha \rightarrow x_1$, $x_1 \rightarrow 1$, $x_2 \rightarrow x_2$, $x_3 \rightarrow x_3$, $x_4 \rightarrow x_4$.

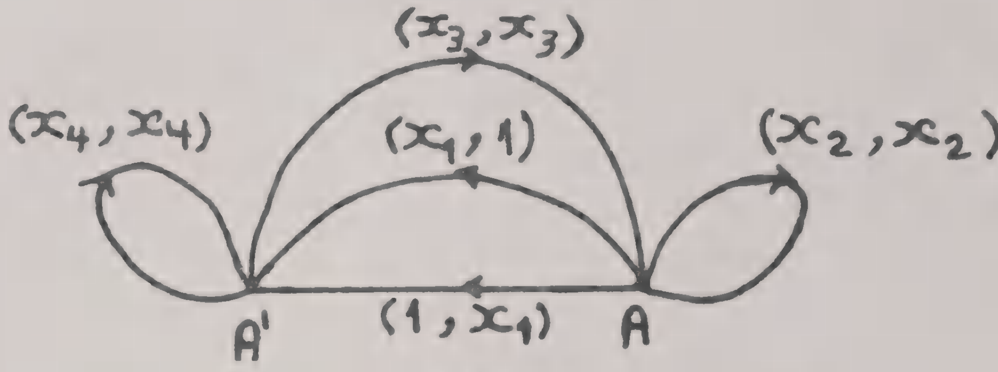


FIG. 2.

3. THE USE OF BASED GRAPHS TO REPRESENT AUTOMORPHISMS

Let Γ be a based graph with two labellings f_1 and f_2 in Fx satisfying the conditions :

1. The labels of any edge are either in X or the identity, and every $x \in X$, is the first label of some edge, and the second label of some edge.

2. T_1 and T_2 are spanning, where T_1 is the set of all edges with first label 1, T_2 is the set of all edges with second label 1

3. No two edges have the same initial vertex, first and second label. (other than (1.1)).

4. Given edges e, e' both having labels $(x, 1)$; $x \in X$, then any path in T_1 from $i(e)$ to $i(e')$ has the same second label as any path in T_1 from $t(e)$ to $t(e')$. Similarly, if e and e' both have labels $(1, x)$; $x \in X$, then any path in T_2 from $i(e)$ to $i(e')$ has the same first label as any path in T_2 from $t(e)$ to $t(e')$.

5. For every $x \in X$, all edges with second label x dave the same second label and similarly, all edges with second label x have the same first label.

6. Any loop in T_1 has second label equal to 1 in Fx, similarly any loop in T_2 has first label equal to 1 in Fx.

Let v_0 be the base point of Γ.

Lemma 3.1—For every reduced word $x_1^{\epsilon_1}, \dots, x_n^{\epsilon_n}$; $\epsilon_i \pm 1$, $i = 1, \dots, n$ there is a reduced path p such that $i(p) = v_0$, $t(p) = v_0$ and $f_1(p) = x_1^{\epsilon_1} \dots x_n^{\epsilon_n}$.

Lemma 3.2—All paths with first label $x_1^{\epsilon_1} \dots x_n^{\epsilon_n} \in Fx$ and $i(p) = v_0 = t(p)$ have the same second label.

From Lemmas 3.1 and 3.2,, we can define a map $\phi : Fx \to Fx$ on reduced words in X by :

$$\phi\,(x_1^{\epsilon_1} \ldots x_n^{\epsilon_n})\,;\ \epsilon_i = \pm 1,\ i = 1, \ldots, n;$$

equals the reduced word equivalent to the second label of any path p ($i\,(p) = t\,(p) = v_0$) with first label $x_1^{\epsilon_1} \ldots, x_n^{\epsilon_n}$.

Similarly, we can define $\psi : Fx \to F_X$ by

$$\psi\,(x_1^{\epsilon_1} \ldots x_n^{\epsilon_n}\,);\ \epsilon_i = \pm 1,\ i = 1, \ldots, n;$$

equals the reduced word equivalent to the first label of any path p ($i\,(p) = t\,(p) = v_0$) with second label $x_1^{\epsilon_1} \ldots x_n^{\epsilon_n}$.

Lemma 3.3—ϕ is a homomorphism.

Theorem 3.1—ϕ and ψ are inverses of each other. and hence they are automorphisms.

Definition 3.1—If (Γ, v_0) is a based graph satisfying the conditions 1 —6, then (Γ, v_0) is called a based graph representing ϕ.

Corollary 3.1—Every automorphism is represented by a based graph (satisfying 1—6).

Definition 3.2—(1) If v is a vertex in (Γ, v_0) with valency 2, $v \neq v_0$ and if (i) v is an extreme vertex of both T_1 and T_2; (ii) the labels $(x, 1)$ and $(1, y)$ of the two edges in $\text{Star}_\Gamma\,(v) = \{e \in (\Gamma, v_0) / i\,(e) = v\}$ occur as the labels of some other edges of (Γ, v_0), then v and its adjacent edges are redundant.

(2) If $v = v_0$ satisfies condition (i) but not (ii) then we say that v and its adjacent edges are singular.

Definition 3.3—From a based graph, we get a smaller graph as follows :

M1. Identify vertices by edge collapses and edge folds, collapsing all edges labelled $(1, 1)$ and folding all edges with both labels the same and with the same initial vertex, until no further such folds are possible.

M2. Remove, all the redundant vertices, and their adjacent edges.

M3. After removing all redundant vertices, combine every two edges e_1 and e_2 say, where e_1 is an edge labelled $(x, 1) \in T_2$ with $i\,(e_1) = u$, $t\,(e_1) = v$ and e_2 is an edge labelled $(1, y) \in T_1$ with $i\,(e_2) = v$, $t\,(e_2) = u_1$ into one edge e labelled (x, y) with $i\,(e) = u$, $t\,(e) = u_1$ whenever v is singular.

Observation : The operation M2 and M3 do not introduce any edge to which M1 can be applied or any redundant vertices.

Definition 3.4—A based graph with,

(i) No edge labelled (1, 1),

(ii) No possible edge fold (for edges with same labels),

(iii) No redundant vertices,

(iv) No singular vertices,

is called a minimal based graph.

Proposition 3.1—After applying M1 to the graph (Γ, v) T_1 and T_2 are disjoint and spanning; every vertex v with valency 2 is either singular or redundant; if v is singular then neither of the labels $(x, 1)$ and $(1, y)$ of the two edges in $\mathrm{Star}_\Gamma(v)$ occurs as the label of any other edges of (Γ, v_0).

Theorem 3.2—After applying M1, M2 and M3 to a based graph Γ. We still have a based graph.

Theorem 3.3—In a minimal based graph the following condition holds (2′) T_1 and T_2 are disjoint maximal trees.

4. The main Result

Definition 4.1[2,3,9,10]—If Γ is a graph and $v \in V(\Gamma)$ let $\mathrm{Star}_\Gamma(v) = (e \in \Gamma / i(e) = v)$ A morphism $f : \Gamma \to \Gamma'$ of graphs induces a map

$f_v : \mathrm{Star}_\Gamma(v) \to \mathrm{Star}_{\Gamma'}(f(v))$. We say f is an immersion if f_v is injective for each $v \in V(\Gamma)$.

Definition 4.2[2]—A morphism $f : \Gamma \to \Gamma'$ of finite graphs is called special if it is a composition of edge collapses and edge folds.

Lemma 4.1[1]—Let Y be a finite graph with one vertex y_0 labelled $1 \in Fx$, and one edge labelled $x^{\pm 1}$ for each $x \in X$. Let l denote the labelling of $Y \to Fx$. For each labelling f of $E(\Gamma) \to F_X$ satisfying (1) (page 7) extend f by mapping every vertex to $l \in Fx$. Then there exists a unique graph morphism $\tilde{f}$ of $\Gamma \to Y$ such that $l.\tilde{f} = f$.

Theorem 4.1[1]—Let (Γ, v_0) be a based graph satisfying (2′) and let $\tilde{f_1}$ and $\tilde{f_2}$ be defined, and let ϕ be the automorphism of Fx represented by Γ. Then the following conditions Gersten[2] holds :

G1. T_1 and T_2 are maximal trees of Γ and $E(T_1) \cap E(T_2) = \phi$,

G2. $\tilde{f_1}/(\Gamma - E(T_1))$ and $\tilde{f_2}/(\Gamma - E(T_2))$ are both immersions,

G3. $\widetilde{f_1}$ and $\widetilde{f_2}$ are special, $\ker \widetilde{f}_1^* = \ker \widetilde{f}_2^*$ and $\widetilde{f}_2^* \cdot \widetilde{f}_1^{-1*} = \widetilde{\phi}$, where $\widetilde{f}_1^*, \widetilde{f}_2^* : \pi_1(\Gamma, v) \to \pi_1(Y, y_0)$; $\pi_1(\Gamma, v)$ is the set of equivalence classes of paths p such that $i(p) = t(p)\, v$, $\widetilde{\phi}$ is the automorphisms of $\pi_1(Y, y_0)$ corresponding to ϕ.

Now we will prove the converse to Theorem 4.1.

Theorem —A graph satisfying conditions G1—3 of Gersten[2] gives a based graph satisfying (2′).

PROOF : G1 is precisely (2′). Also from G1, since T_1 and T_2 are maximal disjoint trees, then any loop in T_1 at the vertex v has second lebel which equal to 1 in Fx. Otherwise contradicting G1. Similarly, any loop in T_2 has first label which equal to 1 in Fx. Then (6) holds.

From G2, since $f_i/(\Gamma - E(T_i))$; $i = 1, 2$ is an immersion, then f_i is injective on $\mathrm{Star}_\Gamma(v)/T_i$ which means that (3) holds.

Now to prove that (4) and (5) hold, iet (Γ, v_0) be a based graph with two labellings f_1 and f_2, and consider the following assumptions :

(a) T_1 is spanning tree,

(b) $f_1(p) = 1 \Rightarrow f_2(p) = 1$ whenever p is a circuit at v_0 Hoare,

(c) $f_2(p)$ is reduced whenever p is a reduced path with no edges in T_2 Hoare,

(d) $T_1 \cap T_2 = \phi$.

Then we prove that (Γ, v_0) as the following property: G′4. If e and e' are two distinct edges with $f_1(e) = f_2(e') = 1$, then $e, e' \in T_2$. Moreover T_1 is a tree and if α and β are the unique paths in T_1 from $i(e)$ to $i(e')$ and $t(e)$ to $t(e')$ respectively, then $f_2(\alpha) = f_2(\beta)$. To do so, if T_1 contains a non-trivial circuit p then using (a), T_1 contains a non-trivial reduced circuit p at v_0. From assumption (d), p has no edges in T_2, and on using (c) and (b), $f_2(p)$ is reduced with $f_2(p) = 1$, so p is empty, giving a contradiction, therefore T_1 must be a tree.

Now suppose that $e \in T_2$, and let p be the circuit $q.\, e'\, \bar{\beta}.\bar{e}.\alpha.\, \bar{q}$ as in Fig. 3, where q is the unique path in T_1 from v_0 to $i(e')$ (since T_1 spanning tree). Then $f_1(p) = f_1(e')\, f_1(e) = 1$, and on using assumption (b),

$$f_2(p) = f_2(q)\, f_2(e')\, f_2(\bar{\beta})\, f_2(\bar{e})\, f_2(\alpha)\, f_2(\bar{q}) = 1. \qquad \text{...(1)}$$

According to (d) $\bar{\beta}.\bar{e}.\alpha$ has no edges in T_2, therefore by (c) $f_2(\bar{\beta}.\bar{e}.\alpha)$ is reduced and

Fig. (3)

from (1) $f_2\,(\overline{\beta.\bar{e}.\alpha}) = f_2\,(e)^{-1}$. Thus α and β are empty and $f_2\ (e) = f_2\,(e') = 1$. Then $e'.\,e$ is a reduced path which contradicts (c). Thus $e \in T_2$. Similarly $e' \in T_2$, and since $f_2\,(\beta)$ and $f_2\,(\alpha)$ are reduced according to (c) with $f_2\ (e'.\ (e'.\overline{\beta}.\bar{e}.\alpha) = 1$, then $f_2\,(\beta) = f_2\,(\alpha)$.

Similarly, on taking the assumptions;

(a′) T_2 is spanning tree,

(b′) $f_2\,(p) = 1 \Rightarrow f_1\,(p) = 1$, whenever p is a circuit at v_0,

(c′) $f_1\,(p)$ is reduced whenever p is a reduced path with no edges in T_1,

(d′) $T_1 \cap T_2 = \phi$;

and applying the same arguments as in the proof of G′4, one can prove that $(\Gamma,\, v_0)$ has the following property; if e and e' are two distinct edges with $f_2\,(e) = f_2\,(e') = 1$, then $e, e' \in T_1$, moreover T_2 is a tree and if α and β are the unique paths in T_2 from $i\ (e)$ to $i\ (e')$ and $t\ (e)$ to $t\ (e')$ respectively, then $f_1\,(\alpha) = f_1\,(\beta)$. This completes the proof of the theorem.

REEERENCES

1. A. G. A. E. Abdel-Gawad, Ph. D. Thesis, Birmingham University, 1986.
2. S. M. Gersten, Fixed points of Automorphisms of Free Groups, *Advances Math.* (to appear).
3. S. M. Gersten, *Inv. Math.* 71 (1983), 567-91.
4. M. Jr. Hall, The MacMillan Company, 1959.
5. A. H. M. Hoare, *Can. J. Math.* 31 (1979), 112-33.
6. A. H. M. Hoare, Automorphisms of free groups. (to appear).
7. R. C. Lyndon and P. E. Schupp, *Combinatorial Group Theory*. Vol. 89 *Ergebnisse*, Berlin Springer, 1977.
8. J. P. Serre, *Trees*. Springer-Verlag, Berlin 1980.
9. J. R. Stallings, *Inv. Math.* 71 (1983), 551-65.
10. J. R. Stallings, *Proceding of Alto Conference on Cambinatorial group theory and very low-dimensional Topology*. (*July* 1984).

Indian J. pure appl. Math., 20 (1) : 49-57, January 1989

ARTINIAN (NOETHERIAN) PART OF A GOLDIE RING

KHANINDRA CHANDRA CHOWDHURY

Department of Mathematics, Gauhati University, Guwahati 781014, *Assam*

(*Received* 4 *August* 1987; *after revision* 1 *January* 1988)

An Artinian ring with unity (a Noetherian ring) is always a Goldie ring. But the converse in not true. Here we discuss what can be thought as an Artinian (Noetherian) part of a Goldie ring and get some interesting results, mainly on right (Left) Artinian radical (i.e. the sum of all the right (left) ideals which are Artinian as right (left) R-modules) of a Goldie ring R.

1. INTRODUCTION

We know that an Artinian ring (with unity) is a Noetherian ring and a Noetherian ring is always a Goldie ring. But the converse is not true. For example the commutative integral domain e.g. $Z[X_i, i = 1, 2, ...]$, where $X_iX_j = X_jX_i$, is a Goldie ring but it is not Noetherian. However a Goldie ring may have a part which behaves as an Artinian (Noetherian) ring or an Artinian (Noetherian) module. Here, unless otherwise specified, a ring R will be a ring with unity.

A ring R satisfies 'right essential descending chain condition (r.e.d.c.c.)' if any descending chain of right essential ideals stops after a finite number of steps. Similarly the 'left essential descending chain condition (l.e.d.c.c)' may be defined. All finite rings and right (left) Artinian rings are rings with r.e.d c.c. (l.e.d.c.c.).

The sum $A(R)$ of all the Artinian right ideals of R is the 'right Artinian radical' of R (An Artinian right ideal means a right ideal of R which is Artinian as a right R-module). We have $A(R) = R$ if and only if R is a right Artinian ring and $A(R) = 0$ if and only if R has no Artinian right ideal. Similarly the 'left Artinian radical' $B(R)$ of R is defined. $A(R)$ and $B(R)$ are ideals of R. When $A(R) = B(R)$ $(= A)$ and A is Artinian ideal then A is called the 'Artinian radical of R. If R is left and right Noetherian then $A(R) = B(R)$ (page 59, Chatters and Hazarnavis[2]).

A right ((left), two sided) ideal of R is a 'countable right' (left, two sided) ideal if it is countable as a set. All the rings which are countable as sets have countable ideals. For example, let

$$F = Z/2Z \text{ and } R = \begin{bmatrix} Z & F \\ O & Z \end{bmatrix}.$$

Here R is left and right Noetherian and N is the set of strictly upper triangular elements of R. Then $A(R) = B(R) = N$ which is countable.

He we prove that if a minimal ideal I of a semiprime right Goldie ring is Artinian as a left R-module then it is a right module over a simple Artinian extension of an epimorphic image of R. Another result we prove here is : if R is a semiprime right Goldie ring satisfying the r.e.d.c.c. and a minimal ideal I of R is Noetherian as a right R-module then I is Artinian as a right R-module.

The ring $R = \begin{bmatrix} Z & Q \\ O & Q \end{bmatrix}$ is right Noetherian (Hence it is right Goldie). The right ideals $e_{12}\, Q$, $e_{22}\, Q$ of R are Artinian (in fact minimal). Now it follows that $A(R) = e_{12}\, Q + e_{22}\, Q$ and $A(R)$ is a minimal ideal of R (Moreover it is countable). Here we prove that in a semiprime right and left Goldie ring satisfying the r.e.d.c.c. and l.e.d.c.c. if $A(R)$ and $B(R)$ are minimal ideals and $A(R)$ is Noetherian as an R-module then $A(R) = B(R)$.

Other two interesting results we prove here are :

(1) In a semiprime fully left Goldie ring R if its minimal prime ideals are Noetherian as R-modules, then in some special cases, $A(R) + r(A(R))$ contains a regular element.

(2) If R is a semiprime fully left and fully right Goldie ring such that its minimal prime ideals are Noetherian as R-modules and its Artinian radical A is minimal and countable then A is a direct summand of R.

2. Preliminaries

A right (left) Goldie ring R is 'fully right (left) Goldie' if any homomorphic image of R is right (left) Goldie.

2.1. We now prove :

Lemma 2.1.1 — If I is a right ideal of a semiprime ring R where $R/r(I)$ is Artinian ring, then I is Artinian as a right R-module.

Proof : Since $R/r(I)$ is a right Artinian ring, it is Artinian as a right $R/r(I)$—module. So the submodule $I/r(I)$ is Artinian as a right $R/r(I)$—module.

If $P_1 \supseteq P_2 \supset \ldots$, is a descending chain of R-submodules of I, then $Q_1 \supseteq Q_2 \supseteq \ldots$ is a descending chain of $R/r(I)$ submodules of $I/r(I)$, where

$$Q_i = \{p_i + r(I) \mid p_i \in P_i\}. \text{ So } Q_t = Q_{t+1} = , \ldots, \text{ for some } t \in Z^+$$

R being semiprime, for any right ideal J $(\subseteq I)$ we get $J \cap r(I) = 0$. And hence $P_t = P_{t+1} = \ldots$ Thus I is Artinian as a right R-module.

Lemma 2.1.2 — A semiprime right (left) Goldie ring R satisfying the r.e.d.c.c. (l.e.d.c.c.) is a quotient ring.

PROOF : Since R is semiprime right Goldie, it has a regular element, (say) $d \in R$ [Theorem 1.10 of Chatters and Hazarnavis[2]]. Then each of $dR, d^2R, d^3R, \ldots$ is right essential ideal of R (Lemma 1.11 of Chatters and Hazarnavis[2]) and R being with r.e.d.c.c. the descending chain $dR \supseteq d^2R \supseteq d^3R \supseteq \ldots$ gives $d^tR = d^{t+1}R$, for some $t \in Z^+$. Hence $R = dR$ which gives $d^{-1} \in R$. Thus R is a quotient ring. (Similarly for left Goldie ring).

If P is an ideal of R, then from the natural epimorphism $\nu : R \to R/P$ we get that for any right (left) essential ideal X of R/P, $\nu^{-1}(X) = I$ is a right (left) essential ideal of R (Goodearl[3], Proposition 1.1). [It is to be noted that a finite integral domain is a field follows from this Lemma easily]. Now we get

Lemma 2.1.3 — If R is a ring satisfying the r.e.d.c.c. (l.e.d.c.c.) and P an ideal of R, then the ring R/P also satisfies the r.e.d.c.c. (l.e.d.c.c.).

From Lemma 2.1.2 and from the Lemma 12, Appendix B of Jacobson[4] we get :

Lemma 2.1.4 — If R is a semiprime right Goldie ring satisfying the r.e.d.c.c. and P is any annihilator prime ideal of R, then R/P is a quotient ring.

Now we prove the following :

Lemma 2.1.5 — For any minimal ideal I of a semiprime right Goldie ring R if $P = r(I)$ then

$$Z(I) = \{x \in I \mid xc = 0 \text{ for some regular } c + P \in R/P\}$$
$$= 0.$$

PROOF : For $x, y \in Z(I)$ we get $c + P$, $d + P$ regular in R/P such that $xc=0$, $yd = 0$. And I being minimal, R/P is prime Goldie. So by Proposition 1.11 of Chatters and Hazarnavis[2], $(c + P)\, R/P$ and $(d + P)\, R/P$ are right essential ideals of R/P and so is their intersection $(c + P)\, R/P \cap (d + P)\, R/P$. Therefore it contains a regular element, say $e + P$. Thus $e + P = (c + P)(u + P) = (d + P)(v + P)$ for some $u, v \in R$, which gives $e - cu,\ e - dv \in P\ (= r(I))$. So $I(e-cu) = I(e-dv) = 0$, or, $x(e - cu) = y(e - dv) = 0$, for $x, y \in I$. Therefore $xe = ye = 0$ (since $xc = yd = 0$) which gives $(x - y)\,e = 0$. Since $e+P$ is regular in R/P, $x-y \in Z(I)$. Again for $r \in R$,

$$(r + P)^{-1}((c + P)\, R/P) = \{u + P \mid (r + P)(u + P) \in (c + P)\, R/P\}$$
$$= \{u + P \mid ru + P \in (c + P)\, R/P\}$$

is a right essential ideal of R/P (Goodearl[3], Proposition 1.19d). Therefore it contains a regular element, say $f + P$. Now $rf + P \in (c + P)\, R/p$ gives $rf - cv \in P$ for some $v \in R$. So $I(rf - cv) = 0$ which gives $xrf = 0$ (since $xc = 0$) or $xr \in Z(I)$. Thus

$Z(I)$ is an ideal of R. Moreover if $x \in Z(I)$, $(x + P)(c + P) = P$ for some regular $c + P$, which implies $x \in P$ or $Ix = 0$. Hence $Z(I)\,x\ (\subseteq Ix) = 0$. It therefore follows that $(Z(I))^2 = 0$. R being semiprime, we finally get $Z(I) = 0$.

Lemma 2.1.6 — If a minimal ideal of a semiprime right Goldie ring R is Artinian as a left R-module, then for any regular $d + P \in R/P$ (where $P = r(I)$) we get $I = Id$.

PROOF : Let $d + P$ be regular in R/P. Since I is Artinian as a left R-module, the chain.

$Id \supseteq Id^2 \supseteq \ldots$ stops after a finite number of steps. Let $Id^t = Id^{t+1}$, for some $t \in Z^+$. So for each $x \in I$ there is some $y \in I$ such that $(x - yd)\,d^t = 0$. Since $d^t + P$ is regular in R/P, it therefore follows that $x - yd \in Z(I)$. But $Z(I) = 0$ (Lemma 2.1.5). Hence $x = yd$ which gives $I \subseteq Id$. Hence $I = Id$.

2.2. *Lemma* 2.2.1 — (a) If I is a countable right ideal of a left Goldie ring, then there is a finite subset S of I such that $r(S) = r(I)$.

(b) Let R be a left Goldie ring and I a countable ideal of R. If I is Artinian as a right R-module then $R/r(I)$ is a right Artinian ring.

PROOF : (a) Consider $y_1 \in I$. Then $r\{y_1\} \supseteq r(I)$. Choose $y_2 \in I$, $y_2 \neq y_1$. Then $r\{y_1\} \supseteq r\{y_1, y_2\}$. Thus we get a descending chain of right annihilators,

$$r\{y_1\} \supseteq r\{y_1, y_2\} \supseteq r\{y_1, y_2, y_3\} \supseteq \ldots$$

The ring being left Goldie, the above chain stops after a finite number of steps[2] (page 2). Suppose $r\{y_1, \ldots, y_n\} = r\{y_1, \ldots, y_{n+1}\} = \ldots = r(I)$. Thus $S = \{y_1, \ldots, y_n\}$ such that $r(I) = r(S)$.

(b) From above $r(I) = r(y_1) \cap \ldots \cap r(y_n)$.

Hence by 4.5 (a) of Chatters and Hazarvanis[2], the module $R/r(I)$ is embedded in the module $y_1 R \oplus \ldots \oplus y_n R$ as a right R-module. Since I is Artinian as a right R-module and for all i, $y_i R \subseteq I$ we have that $y_i R$ is Artinian as a right R-module. Hence $R/r(I)$ is Artinian as a right R-module which implies that $R/r(I)$ is Artinian as a right $R/r(I)$ — module. Thus $R/r(I)$ is a right Artinian ring.

Lemma 2.2.2 — If the ideal in Lemma 2.2.1 (b) is a minimal ideal of R then the ring $R/r(I)$ is an Artinian ring.

PROOF : Since I is minimal, $r(I)$ is a prime ideal. Therefore $R/r(I)$ is a prime ring. Hence it is simple (ex. 14 (10 (4)) of Anderson and Fuller[1]). So it is simple left Artinian (Proposition 13.5 of Anderson and Fuller[1]). Thus $R/r(I)$ is (simple), Artinian.

2.3. *Lemma* 2.3.1 — Let R be a semiprime right Goldie ring. Then for an annihilator ideal I containing an ideal U, there is an ideal X such that $IX \subseteq U$.

PROOF : Consider the collection $\{J\}$ of all the annihilator ideals of R contained in I and containing U. Since R is right Goldie, the collection has a maximal element, say I_2 such that $I = I_1 \supseteq I_2 \supseteq U$.

Similarly we get annihilator ideals I_3, I_4, ... etc. such that $I = I_1 \supseteq I_2 \supseteq I_3 \supseteq \dots$ and each of these contains U. Since R is semiprime right Goldie, this descending chain is finite (Corollary 1.15 of Goodreal[3]) and so

$$I = I_1 \supseteq I_2 \supseteq \dots \supset I_{n-1} \supseteq U = I_n.$$

Now if

$$r\,(I_j/I_{j+1}) = P_j/I_{j+1}, \text{ then } I_jP_j \subseteq I_{j+1}. \text{ Set } X = P_1P_2 \dots P_{n-1}.$$

Then,

$$IX = (I_1P_1)\, P_2 \dots P_{n-1} \subseteq (I_2P_2) \dots P_{n-1}.$$

Thus

$$IX \subseteq U.$$

We now use the corresponding results of the above Lemma in the following

Lemma 2.3.2 — Let R be a semiprime fully left Goldie ring and B is an annihilator ideal. If C is an ideal of R then there is an ideal Y of R such that $YB \subseteq BC$.

PROOF : Since $BC \subseteq B$, by Lemma 2.3.1 we get an ideal Y of R such that $YB \subseteq BC$.

From Corollary 4.4 of Chatters and Hazarnavis[2] we see Goldie rings such that if R/C in above is Artinian then R/Y is Artinian.

We now here consider semiprime fully left Goldie ring R where, in above Lemma if R/C is Artinian then R/Y is Artinian.

Moreover as in the same Corollary the Goldie ring R has minimal prime ideals which are Noetherian as R-modules.

Thus we prove :

Lemma 2.3.3 — Let R be a semiprime fully left Goldie ring as above and is such that P is a minimal prime ideal which is Neotherian as an R-module and R/P is Artinian. Then the right Artinian radical $A\,(R)$ of R is not contained in P.

PROOF : Let T be the set of all ideals I of R such that R/I is Artinian. And R being semiprime Goldie it has a finite number of minimal prime ideals each of which is an annihilator ideal, say these are $P_1, \dots, P_k$. So $P = P_i$ for some i. If $P_i \in T$ for all i, then from the monomorphism

$$R \to R/P_1 \oplus \dots \oplus R/P_k \text{ (since } P_1 \dots P_k = 0).$$

We get R as Artinian and hence right Artinian. So $A(R) = R$. Hence $A(R) \nsubseteq P$ (since P is a prime ideal).

Suppose for some j, $P_j \notin T$. If an annihilator ideal B and an ideal C of R be such that $B \notin T$ and $C \in T$, then by Lemma 2.3.2 there is a $D \in T$ such that $DB \subseteq BC$. Thus if $P_2 \in T$, $P_1 \notin T$ then there exists $Q_1 \in T$ such that $Q_1P_1 \subseteq P_1P_2$. Thus we have $Q_1P_1P_3 \ldots P_k \subseteq P_1P_2P_3 \ldots P_k = 0$. In this way P_j $(\notin T)$ can be moved to the right in the product $P_1P_2 \ldots P_k$ until we obtain $XY=0$, where X is a product of elements of T and Y is a product of minimal prime ideals $(\notin T)$.

Let $X = Q_1Q_2 \ldots Q_t$, $Q_i \in T$. Then as in Corollary 4.4 of Chatters and Hazanavis[2] R/X is Artinian. Thus $X \in T$.

Now let $Y = A_1 \ldots A_s$, where each A_i is a minimal prime ideal not contained in T. Since P is prime, $Y \notin P$.

Let $Y \supseteq Y_1 \supseteq Y_2 \supseteq \ldots$ be a descending chain of right ideals. Then this gives a chain $S \supseteq S_1 \supseteq S_2 \supseteq \ldots$ of right ideals of R/X, where $S_i = \{y_i + X \mid y_i \in Y_i\}$. Since $XY = 0$ we get $X \cap Y = 0$. (Since R is semiprime). Thus $X \cap Y_i = 0$, for all i. Since R/X is Artinian we get $S_t = S_{t+1}$ for some $t \in Z^+$. So $Y_t = Y_{t+1}$. Thus Y is Artinian as a right R-module. Therefore $Y \subseteq A(R)$. Hence $A(R) \subseteq P$.

Theorem 3.1 — If a minimal ideal I of a semiprime right Goldie ring R is Artinian as a left R-module, then it is Artinian as a right module over a simple Artinian extension of an epimorphic image of R.

PROOF: Since R/P is prime Goldie, the quotient ring Q of R/P is simple Artinian (Chatters and Hazarnavis[2], Proposition 1.28). By Lemma 2.1.6, since $I = Id$, we get $I + P = (I + P)(d + P)$ or, $I + P = (I + P)(d + P)^{-1}$, for each regular $d + P \in R/P$. It can easily be seen that if $I/P = \{i + P \mid i \in I\}$, then the map $I/P \times Q \rightarrow I/P$

$$(x + P, (r + P)(d + P)^{-1}) \rightarrow (xr + P)(d + P)^{-1}$$

makes I/P a right Q-module.

Now consider the map $I \times Q \rightarrow I$

$$(i, q) \rightarrow X,$$

where the image x is given by the condition

$$(i + P)q = x + P.$$

We first show that x is unique.

If for $x, y \in I$, $x + P = y + P$, then $x - y \in P$ $(= r(I))$.

Therefore $x - y \in I \cap r(I)$.

Since R is semiprime, $I \cap r(I) = 0$ which gives $x = y$.

And if $i, j \in I$ and $q \in Q$ let $(i + P)q = x + P$ and $(j + P)q = y + P$.

Then

$$\begin{aligned}((i + j) + P)q &= ((i + P) + (j + P))q \\ &= (i + P)q + (j + P)q \text{ (since } I/P \text{ is a right } Q\text{-module)} \\ &= (x + P) + (y + P) \\ &= (x + y) + P.\end{aligned}$$

Thus

$$(i + j)q \; (= x + y) = iq + jq.$$

Similarly for $q, 1 \in Q, i \in I$ we get

$$i\,(q1) = (iq)1.$$

So I is a right module over the simple Artinian ring Q which is an extension as a ring of an epimorphic image R/P of R.

Theorem 3.2 — Let R be a semiprime Goldie ring satisfying the r.e.d.c.c. If minimal ideal I of R is Noetherian as a right R-module then I is Artinian as a right R-module.

PROOF : If $P = r\,(I)$ then the ring R/P is prime Goldie and by Proposition 1.28 of Chatters and Hazarnavis[2] the quotient ring Q of R/P is simple Artinian, and the well defined map

$$I \times R/P \to I$$

$$(i, r + P) \to ir$$

makes I a right R/P-module (Since $P = r\,(I)$). By Lemma 2.1.2, $R/P = Q$. Therefore I is a right module over the simple Artinian ring $(Q =)$ R/P. Since I is Noetherian as a right R-module it is Noetherian as a right R/P module. (For any R-submodule of I is an R/P-submodule and vice-versa). Therefore I is finitely generated as a right module over the simple Artinian ring R/P. It follow that I is Artinian as a right R/P-module (Proposition 10.18 of Anderson and Fuller[1]). And hence I is Artinian as a right R-module.

We have already seen Goldie ring whose right Artinian radical is minimal and countable. Thus we consider semiprime fully left Goldie ring R where $A(R)$ is minimal, countable and Artinian as a right R-module. We now prove

Theorem 3.3 — Let R be a semiprime fully left Goldie ring as above and its minimal prime ideals are Noetherian as R-modules. Then $A(R) + r(A(R))$ contains a regular element.

PROOF : Suppose for any minimal prime ideal P, $P \supseteq r\,(A(R))$. Then $P \supseteq A(R) + r(A(R))$. Since $A(R)$ is minimal, countable and Artinian as a right R-module,

and $P \supseteq r(A(R))$ gives $P = r(A(R))$, by Lemma 2.2.2, R/P is Artinian which gives by Lemma 2.3.3, $A(R) \subseteq P$. So $P \supseteq A(R) + r(A(R))$. Thus in any case $P \supseteq A(R) + r(A(R))$.

If S is a semiprime left Goldie ring and I is not contained in any minimal prime ideal of it then for any left ideal K of S with $I \cap K = 0$ we get $I(KS) = 0$ which gives $KS \subseteq P$ (for all P). But the intersection of all such P being zero, we get $K = 0$. Hence I is an essential left ideal of S. Therefore it contains a regular element. Thus $A(R) + r(A(R))$ contains a regular element.

Theorem 3.4 — Let R be a semiprime right and left Goldie ring satisfying the r.e.d.c.c. and l.e.d.c.c.

If both the right Artinian radical $A(R)$ and the left Artinian radical $B(R)$ are minimal ideals and $A(R)$ is Noetherian as a right R-module as well as a left R-module (for a right Noetherian and a left Noetherian ring there are so) then $A(R) = B(R)$ $(= A)$.

Proof : Here $A(R)$ is a minimal ideal of R and is Noetherian as a right R-module. So by Theorem 3.2 $A(R)$ is Artinian as a right R-module. Since $A(R)$ is Noetherian as a left R-module and R is semiprime left Goldie with l.e.d.c.d., we have $A(R)$ is Artinian as a left R-module. So $A(R) \subseteq B(R)$. But $B(R)$ minimal gives us $A(R) = B(R)$.

Theorem 3.5 — Let R be a semiprime fully left Goldie ring with Noetherian minimal primes where the Artinian radical A is minimal and countable. Then A is a direct summand of R.

Proof : By above $A + r(A)$ contains a regular element say, c. Then $c = a + x$, $a \in A$, $x \in r(A)$ and wet get $Ac \supseteq Ac^2 \supseteq \ldots$ Since A is an Artinian left ideal, Ac is also Artinian left ideal. Hence for some $t \in Z^+$, $Ac^t = Ac^{t+1}$. Then for $\alpha \in A$, we get $\beta \in A$ such that $\alpha = \beta c$ (since c is regular). Thus $A \subseteq Ac$. And A being largest Artinian left ideal of R we therefore get $A = Ac$. So $a = ec$, for some $e \in A$. Then $c = a + x = ec + x$ gives $e = e^2$. And for $b \in A$ we get $b =$ be which gives $A \subseteq \mathrm{Re}$. So $A = \mathrm{Re}$. Consider $J = \{x - ex \mid x \in A\}$. Then J is a right ideal and J^2 $(\subseteq AJ = \mathrm{Re}\, J) = 0$. Since R is semiprime, it therefore implies that $J = 0$. Thus $x = ex$ for $x \in A$. Hence $A = eA$ which gives $A \subseteq eR \subseteq A$, i.e. $A = \mathrm{Re}$. Thus $A = Re = eR$. Therefore e is central and so $(1 - e)$ is in centre of R and the ideal $(1 - e)R$ is such that $R = A + (1 - e)R$ and since $er \cap (1 - e)R = 0$, A is a direct summand.

Acknowledgement

The author wishes his gratitude to Professor B.K. Tamuli, Head of the Department of Mathematics (retired), Gauhati University, for his help in this work. He is also thankful to Dr R.N. Gupta of Department of Mathematics, Panjab University, Chandigarh, who helped him with some examples long ago.

References

1. R. W. Anderson and K. R. Fuller, *Rings and Categories of Modules* Springer-Verlag, Berlin, 1973.
2. A. W. Chatters and C.W. Hazarnavis, *Rings with Chain Conditions.* Pitman Advanced Publishing Program, 1980.
3. K.R. Goodearl, *Ring Theory; Nonsingular Rings and modules*, Marcel Dekker Inc., 1976.
4. N. Jacobson, *Structure of Rings*, AMS, 1968.
5. J. Lambek, *Lectures on Rings and Modules*, Blaisdell Publishing Co., 1966.

Indian J. pure appl. Math., 20 (1) : 58-63, January 1989

A GENERALIZATION OF STRONGLY REGULAR NEAR-RINGS

P. DHEENA

Department of Mathematics, Annamalai University
Annamalainagar 608 002

(*Received* 11 *May* 1987; *after revision* 9 *May* 1988)

In this paper we introduce the notion of s-weakly regular near-rings similar to the notion introduced for rings by Gupta[1]. We give some characterizations of s-weakly regular near-rings.

DEFINITIONS AND NOTATIONS

Throughout this paper N stands for a zerosymmetric right near-ring. For any $x \in N$, $<x>$ stands for the principal ideal generated by x. N is said to be s-weakly regular if for each $a \in N$, $a = xa$ for some $x \in <a^2>$.

For any subset S of N, we denote the set $\{n \in N : nS = 0\}$ by $A(S)$. We denote $A(\{a\})$ by $A(a)$. An ideal P of N is called completely prime (semicompletely prime) if $ab \in P$ implies $a \in P$ or $b \in P$ ($a^2 \in P$ implies $a \in P$). An ideal I of N is said to be prime if for any two ideals A, B of N, $AB \subseteq I$ implies either $A \subseteq I$ or $B \subseteq I$. This is equivalent to the following condition. For all finite set $A_1, A_2 \ldots. A_n$ of ideals of N, $A_1 A_2 \ldots. A_n \subseteq I$ implies atleast one $A_i \subseteq I$. An ideal minimal in the set of all prime ideals of N is called minimal prime ideal in N. An ideal I of N is called semiprime if for any ideal J of N, $J^2 \subseteq I$ implies $J \subseteq I$. A subset M of N is called an m-system if for any a, b in M, there exists $a_1 \in <a>$ and $b_1 \in <b>$ such that $a_1 b_1 \in M$. N is said to have IFP (insertion of factors property) if for a, b in N, $ab = 0$ implies $anb = 0$ for all $n \in N$. N is said to be reduced if it is without nonzero nilpotent elements. If N is a zerosymmetric reduced near-ring, then by Pilz[4] (p. 290), N has IFP and for any x, y in N, $xy = 0$ implies $yx = 0$. For the basic terminology and notation we refer to Pilz[4].

SECTION 1

Lemma 1— If N is a reduced near-ring, then for any $0 \neq a \in N$,

(1) $A(a)$ is a semicompletely prime ideal,

(2) $N/A(a)$ is reduced and the residue class $\bar{a}$ of a mod $A(a)$ is a nonzero divisor,

(3) $x_1 x_2 \ldots x_n = 0$ implies $<x_1> <x_2> \ldots <x_n> = 0$

for any $x_1, x_2, \ldots, x_n$ in N.

PROOF : (1) Since N has IFP, $A(a)$ is an ideal. Suppose $x^2 \in A(a)$. Then $0 = x^2 a = x(xa) = xax$, so $(xa)^2 = 0$. Thus $xa = 0$ and hence $A(a)$ is a semicompletely prime ideal.

(2) Since $A(a)$ is a semicompletely prime ideal, $N/A(a)$ is reduced.

Suppose there exists $b \in N$ with $\bar{b}\,\bar{a} = \bar{0}$. Then $\bar{a}\,\bar{b} = \bar{0}$ and hence $ab \in A(a)$. Thus $(ba)^2 = baba = 0$, so $ba = 0$ and hence $\bar{b} = \bar{0}$.

(3) Suppose $x_1 x_2 \ldots x_n = 0$ for some $x_1, x_2, \ldots, x_n$ in N. Since N is reduced, $A(S)$ is an ideal for any subset S of N. Since $x_1 \in A(x_2 \ldots x_n)$, we have $< x_1 > \subseteq A(x_2 \ldots x_n)$ so that $< x_1 > x_2 \ldots x_n = 0$. Hence $x_2 \ldots x_n < x_1 > = 0$. Since $x_2 \in A(x_3 \ldots x_n < x_1 >)$ we have $< x_2 > \subseteq A(x_3 \ldots x_n < x_1 >)$ so that $<x_2> x_3 \ldots x_n < x_1 > = 0$. Hence $x_3 \ldots x_n < x_1 > < x_2 > = 0$. Continuing this process we arrive at (3).

Theorem 1 — Let N be a reduced near-ring. If M a subset of N, is a non-void m-system such that $0 \notin M$, then there exists a completely prime ideal P of N such that $P \cap M = \phi$.

PROOF : Let M' be a maximal m-system relative to the property $M \subseteq M'$ and $0 \notin M$. M' is obtained using Zorn's Lemma and clearly $M \subseteq M'$. By Pilz[4] (Proposition 2.81) there exists a prime ideal $P\,(\neq N)$ such that $P \cap M' = \phi$. Hence $M' \subseteq N \setminus P$. By Pilz[4] (Corollary 2.80) $N \setminus P$ is an m-system. By the maximality of M', $N \setminus P \subseteq M'$. Thus $N \setminus M' = P$. It can be easily verified that P is a minimal prime ideal in N. Now let us show that P is completely prime. Let $\bar{M}$ be the multiplicative subsemigroup of N generated by $N \setminus P$. We claim that $0 \notin \bar{M}$. If not, there exists $x_1, x_2, \ldots, x_n \in N \setminus P$ such that $x_1 x_2 \ldots x_n = 0$. By Lemma 1, $<x_1> <x_2> \ldots <x_n> = 0 \subseteq P$. Thus $<x_i> \subseteq P$ for some i. Hence $x_i \in P$ which contradicts our assumption. Let $K = \{J : J$ is an ideal of N such that $J \cap \bar{M} = \phi\}$. K is not empty. By Zorn's Lemma K contains a maximal element, say Q. Now we claim Q is prime. For, if there exists ideals A, B such that $Q \subset A$ and $Q \subset B$ then take $a \in A \cap \bar{M}$ and $b \in B \cap \bar{M}$. Then $ab \in \bar{M}$ and $ab \in < a > < b > \subseteq AB$. Hence $< A\ B > \cap \bar{M} \neq \phi$ so that $< A\ B > \nsubseteq Q$ and $A\ B \nsubseteq Q$. Thus Pilz[4] (Proposition 2.61) Q is prime. Now $Q \subseteq N \setminus \bar{M} \subseteq P$. Since P is a minimal prime ideal $Q = N \setminus \bar{M} = P$. Since $\bar{M}$ is a semi-group, P is completely prime.

Corollary 1 — Let N be a reduced near-ring. If M is a non-void multiplicative subsemigroup of N such that $0 \notin M$, then there exists a completely prime ideal P of N such that $P \cap M = \phi$.

Proof : Since every multiplicative subsemigroup is an m-system, the corollary follows immediately.

Now we prove our main theorem.

Theorem 2— The following are equivalent for a near-ring N with identity :

(1) N is s-weakly regular,

(2) N is reduced and every proper prime ideal is maximal,

(3) N is reduced and every proper completely prime ideal is maximal.

Proof : (1) $\Rightarrow$ (2)— Suppose $a \in N$ such that $a^2 = 0$. We have $a = xa$ for some $x \in < a^2 > = 0$, so that $a = 0$. Thus $a^2 = 0$ implies $a = 0$ for every a in N. Hence N is reduced. Let P be a proper prime ideal and suppose P is properly contained in a maximal ideal M. Let $x \in M \setminus P$. Then $x = yx$ for some $y \in < x^2 >$. Now for any $n \in N$, $nx = nyx$ so that $(n - ny)\ x = 0$. By Lemma 1, $< (n - ny) > < x > = 0 \subseteq P$. Since P is a prime ideal and $x \notin P$, we have $n - ny \in P \subset M$. Further $y \in < x^2 > \subseteq M$ whence $ny \in M$ so that $n \in M$. Hence $M = N$, a contradiction.

(2) $\Rightarrow$ (3) is obvious.

(3) $\Rightarrow$ (1)— Let $0 \neq a \in N$. By Lemma 1, $\bar{N} = N/A(a)$ is reduced and $\bar{a}$ is not a zero divisor. Also every proper completely prime ideal of $\bar{N}$ is a maximal ideal in $\bar{N}$. Now let M be the multiplicative semigroup generated by all elements $\bar{a} - \bar{x}\,\bar{a}$ where $x \in < a^2 >$. We claim that $\bar{O} \in M$. If not, by Corollary 1, there exists a completely prime ideal $\bar{P}$ with $\bar{P} \cap M = \phi$. Suppose $< \bar{a}^2 > \subseteq \bar{P}$, then $\bar{a}^2 \in \bar{P}$ whence $\bar{a} \in \bar{P}$. Now for any $x \in < a^2 >$, $\bar{a} - \bar{x}\ \bar{a} \in \bar{P} \cap M$, a contradiction. Suppose $< \bar{a}^2 > \nsubseteq \bar{P}$. Since $\bar{P}$ is maximal, $\bar{P} + < \bar{a}^2 > = \bar{N}$. Hence $\bar{1} = \bar{\alpha} + \bar{x}$ for some $\bar{\alpha} \in \bar{P}$ and for some $\bar{x} \in < \bar{a}^2 >$. So $\bar{1} - \bar{x} = \bar{\alpha}$ and hence $\bar{a} - \bar{x}\bar{a} = \bar{\alpha}\bar{a} \in \bar{P} \subset M$, a contradiction, Thus $\bar{O} \in M$. Now $\bar{O} = (\bar{a} - \bar{x}_1\,\bar{a}) \ldots (\bar{a} - \bar{x}_n\,\bar{a})$ where $x_i \in < a^2 >$. Since $\bar{N}$ is reduced and $\bar{a}$ is not a zero divisor, we have $(\bar{1} - \bar{x}_1)\ (\bar{1} - \bar{x}_2) \ldots (\bar{1} - \bar{x}_n) = \bar{O}$. It can be verified that $\bar{1} = \bar{x}$ for some $x \in < a^2 >$. Hence $(l - x) \in A(a)$. Thus $a = xa$ for some $x \in < a^2 >$.

Corollary 2 (Gupta[1], Theorem 5)— The following are equivalent for a ring A with identity :

(1) A is s-weakly regular,

(2) A is reduced and every proper prime ideal is maximal,

(3) A is reduced and every proper completely prime ideal is maximal.

Proof : When the near-ring N is also a ring, our definition of s-weakly regular coincide with the same given by Gupta[1], since $a = xa$ if and only if $a = ax$ for some $x \in < a^2 >$.

Recall that a near-ring N is said to be strongly regular if given $a \in N$ there exists $x \in N$ such that $a = xa^2$.

Remark 1 : If N is strongly regular, then N is s-weakly regular.

PROOF : Suppose N is strongly regular. Let $a \in N$. Then $a = xa^2$ for some $x \in N$. By Reddy and Murty[5], (Theorem 3), $a = axa$ and $ax = xa$. Thus $a = (ax)axa = xa^2\, xa = ya$ where $y = xa^2\, x \in \langle a^2 \rangle$.

The following example shows that the existence of s-weakly regular rings that are not strongly regular.

Example 1— Consider the ring R given in Example 1 of Gupta[1].

Corollary 3— If N is a strongly regular near-ring, then every proper prime ideal is maximal.

The above corollary follows from Remark 1 and Theorem 2 in which (1) implies (2) is true even without the assumption N contains an identity. The above corollary improves the result of Mason[3] (Lemma 4) namely, if N is a strongly regular near-ring with identity then every proper prime ideal is maximal.

Recall that a near-ring N is called strict weakly regular (weakly regular), if $A^2 = A$ for every N-subgroup (ideal) A of N.

Theorem 3— If N is a reduced strict weakly regular near-ring with identity, then N is s-weakly regular.

PROOF : Suppose N is a reduced strict weakly regular near-ring. Let P be a proper completely prime ideal of N. Then $\bar{N} = N/P$ is strict weakly regular and $\bar{N}$ is without zero divisors. By Jat and Choudhary[2] (p. 179) $\bar{N}$ is simple so that P is maximal. By Theorem 2, N is s-weakly regular.

Lemma 2— If N is s-weakly regular, then every ideal of N is semicompletely prime.

PROOF : Suppose N is s-weakly regular. Let I be an ideal of N and let $a^2 \in I$. We have $a = xa$ for some $x \in \langle a^2 \rangle \subseteq I$. Thus $a \in I$.

Theorem 4— If N is an s-weakly regular near-ring, then every ideal I of N is the intersection of all maximal ideals containing I.

PROOF : Let I be an ideal of N. By Lemma 2, I is semi-completely prime and hence I is semiprime. By Sambasiva Rao[6] (Corollary 2.3) I is the intersection of all prime ideals containing I. By Theorem 2, I is the intersection of all maximal ideals containing I.

Theorem 5— Let N be s-weakly regular. Then

(1) N is weakly regular.

(2) N is a subdirect product of simple reduced near-rings.

Proof : (1) Let I be an ideal of N and $a \in I$. Since $a = xa$ for some $x \in \langle a^2 \rangle \subseteq I$, we have $I \subseteq I^2$. Thus N is weakly regular.

(2) By Theorem 2, N is reduced. By Theorem 4, $\{O\}$ is the intersection of maximal ideals. Hence by Pilz[4] (p. 25) N is isomorphic to subdirect product of simple reduced near-ring.

Section 2

Now we find conditions for an s-weakly regular near-ring to be strongly regular.

Lemma 3— Let N be a reduced near-ring. For any a, b in N, if e is an idempotent in N, then $abe = aeb$.

Proof : Since N is reduced, N has IFP and $xy = 0$ implies $yx = 0$ for any x, y in N. Let e be an idempotent in N and $a, b \in N$. Since $(a-ae)e = 0$, we have $(a-ae)be = 0$. Hence $abe = aebe$. Since $(eb-ebe)e = 0$, we have $eb\,(eb-ebe) = 0$ and $ebe\,(eb-ebe) = 0$ so that $(eb-ebe)^2 = 0$. Hence $eb = ebe$. Thus $abe = aeb$.

Theorem 6— Let N be an s-weakly regular near-ring. N is strongly regular if and only if every N-subgroup of N is an ideal.

Proof : Suppose every N-subgroup of N is an ideal. Let $a \in N$ and $a \neq xa^2$ for all $x \in N$. Since Na^2 is an N-subgroup by Theorem 4, Na^2 is the intersection of all maximal ideals containing Na^2. Hence there exists a maximal ideal M, containing Na^2 such that $a \notin M$. But $a^3 \in M$. By Lemma 2, M is semicompletely prime. Hence $a \in M$, a contradiction. Therefore N is strongly regular.

Conversely let N be strongly regular. Since N is reduced, N has IFP. Let $0 \neq a \in N$. By Reddy and Murty[5], $a = axa$ for some $x \in N$. Let $xa = e$. Then e is an idempotent and $Na = Ne$. Denoting the set $\{n-ne : n \in N\}$ by S, we claim that $A\,(S) = Ne$. Since $(n-ne)\,e = 0$ for any $n \in N$, using IFP, $(n-ne)\,Ne = 0$ so that $Ne\,(n-ne) = 0$. Hence $Ne \subseteq A\,(S)$. Suppose $z \in A(S)$. There exists some $y \in N$ such that $z = yz^2$, Now $(yz-yze)z = 0$. Hence $z = yzez$. By Lemma 3, $z = yzze = ze$. Thus $A\,(S) \subseteq Ne$. Hence $Na = Ne = A(S)$. Since N has IFP, $A(S)$ is an ideal and hence Na is an ideal of N. If Δ is an N-subgroup, then $\Delta = \sum_{a \in \Delta} Na$. Since the sum of any family of ideals is an ideal, Δ is an ideal.

Corollary 4— (Mason[3], Theorem 1)— If N is strongly regular with identity, then every N-subgroup of N is an ideal.

For any subset A of N, we write $\sqrt{A} = \{x \in N : x^k \in A$ for some positive integer $K\}$. Recall that N is said to be regular if given $a \in N$, there exists $x \in N$ such that $a = axa$.

Theorem 7— Let N be s-weakly regular. Then the following are equivalent :

(1) N is regular,

(2) $A = \sqrt{A}$ for every N-subgroup A of N,

(3) N is strongly regular.

PROOF : (1) $\Rightarrow$ (2) — Let A be an N-subgroup of N. Suppose $a \in \sqrt{A}$. Then $a^k \in A$ for some positive integer k. Since N is s-weakly regular $a = xa$ for some $x \in \langle a^2 \rangle$. Since N is regular $a = aya$ for some $y \in N$. Since ya is an idempotent, by Lemma 3, $a = xa = xa(ya) = xya^2 = xy(xya^2)\ a = xyxya^3 = \ldots = za^k$ for some $z \in N$. Thus $a = za^k \in A$. Hence $A = \sqrt{A}$.

(2) $\Rightarrow$ (3)— Let $0 \neq a \in N$ Now $a^3 \in Na^2$ so that $a \in \sqrt{Na^2} = Na^2$. Thus N is strongly regular.

(3) $\Rightarrow$ (1) follows from Theorem 3 of Reddy and Murty[5].

REFERENCES

1. V. Gupta, *Acta Math. Hung.* 43 (1984), 57–61.
2. J. L. Jat and S. C. Choudhary, *Math. Student* 4 (2) (1978), 175–82.
3. G. Mason, *Proc. Edinburgh Math. Soc.* 23 (1980), 27–35.
4. G. Pilz, Near-rings. North-Holland, Amsterdam, 1977.
5. Y. V. Reddy and C. V. L. N. Murty, *Proc. Edinburgh Math. Soc.* 27 (1984), 61–64.
6. V. Sambasiva Rao, *J. Aust. Math. Soc.* 32 (1984), 212–14.

Indian J. pure appl. Math., 20 (1) : 64–74, January 1989

LACUNARY DISTRIBUTION OF SEQUENCES

G. DAS

Department of Mathematics, Utkal University, Vani Vihar, Bhubaneswar

AND

B. K. PATEL

Department of Mathematics, Khallikote College, Berhampur (Ganjam) 760001
Orissa

(*Received* 1 *January* 1988; *after revision* 22 *July* 1988)

In this paper we considered a sequence of known spaces and its inclusion which enable us to define a new type of distribution of sequence called lacunary distribution.

§1. Let $\theta = (K_r)$ be the sequence of positive integers such that

(i) $K_0 = 0$, and $0 < K_r < K_{r+1}$

(ii) $h_r = (K_r - K_{r-1}) \to \infty$, as $r \to \infty$.

Then θ is called a lacunary sequence. The intervals determined by θ are denoted by $I = (K_{r-1}\ K_r]$. The ratio K_r/K_{r-1} will be denoted by q_r.

We use the following known spaces of sequences of real numbers.

$C_1 = \{x = (x_k)$: there exists 'L' such that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (x_k - L) = 0\}.$$

$|\, C_1 \,| = \{x = (x_k)$: there exists 'L' such that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |\, x_k - L \,| = 0\}.$$

$C_\theta = \{x = (x_k)$: there exists 'L' such that

$$\lim_{r\to\infty} \frac{1}{h_r} \sum_{K\in I_r} (x_k - L) = 0\}.$$

$| C_\theta | = \{x = (x_k)$: There exists 'L' such that

$$\lim_{r\to\infty} \frac{1}{h_r} \sum_{K \in I_r} | x_k - L | = 0\}.$$

$AC = \{x = (x_k)$: There exists L such that, uniformly in $i \geqslant 0$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{K=0}^{n-1} (x_{k+1} - L) = 0\}.$$

$| AC | = \{x = (x_k)$: There exists 'L' such that, uniformly in $i \geqslant 0$.

$$\lim_{n\to\infty} \frac{1}{n} \sum_{K=0}^{n-1} | x_{k+1} - L = 0.\}$$

$AC_\theta = \{x = (x_k)$: There exists 'L' such that, uniformly in $i \geqslant 0$.

$$\lim_{r\to\infty} \frac{1}{h_r} \sum_{K \in I_r} (x_{k+i} - L) = 0\}$$

$| AC_\theta | = \{x = (x_k)$: There exists 'L' such that, uniformly in $i \geqslant 0$

$$\lim_{r\to\infty} \frac{1}{h_r} \sum_{K \in I_r} | x_{k+i} - L | = 0\}.$$

It is evident that

$$| C_1 | \subset C_1, \ | C_\theta | \subset C_\theta, \ | AC | \subset AC \text{ and } | AC_\theta | \subset AC_\theta.$$

The space C_1 is the well known $(c, 1)$ summable space, $| C_1 |$ is the strong summble $(c, 1)$ space. The spaces C_θ and $| C_\theta |$ spaces of lacunary and lacunary strongly convergence have been recently introduced by Freedman *et al.*[2]. The well known space AC the space of all almost convergent sequences and defined by Lorentz[3] and $| AC |$ the space of strongly convergence has been recently introduced by Maddox[4-6] and also independently by Freedman *et al.*[2]. The space AC_θ and $| AC_\theta |$ the space of lacunary almost convergence and space of lacunary strongly almostly convergence has been studied by Das and Misra[1].

We record some known results :

Theorem A^2—(a) $| C_1 | \subset C_\theta$, if and only if

$$\underline{\lim}_{r}\, q_r > 1; \qquad ...(1.1)$$

(b) $| C_\theta | \subset C_1$, if and only if

$$\overline{\lim}_{r}\, q_r < \infty \qquad ...(1.2)$$

(c) $| AC | \subset C_\theta |$, for every θ;

(d) $C_1 \subset C_\theta$, if and only if

$$\underline{\lim}_{r} \; q_r > 1; \qquad \ldots(1.3)$$

(e) $C_\theta \neq C_1$, for every θ;

(f) $C_\theta \cap l_\infty \subset C_1$, if and only if

$$\lim_{r} \; q_r = 1; \qquad \ldots(1.4)$$

where l_∞ is the set of all bounded sequences.

(g) $AC \subset C_\theta \cap l_\infty$, for every θ.

§2. In the sequel we have the following results :

Theorem 1— $| AC_\theta | \Leftrightarrow | AC |$ for every 'θ'.

We need the following Lemma for the proof of the theorem.

Lemma 1—Suppose, for given $\epsilon > 0$, $\exists\ n_0$ and p_0 such that

$$\frac{1}{n} \sum_{j=p}^{p+n-1} | x_j - L | < \epsilon \qquad \ldots(2.1)$$

for all $n \geqslant n_0$, $p \geqslant p_0$. Then $\{x_j\} \in | AC |$.

PROOF : Let $\epsilon > 0$ be given. Choose n_0^1, p_0 such that

$$\frac{1}{n} \sum_{j=p}^{p+n-1} | x_j - L | < \epsilon/2 \qquad \ldots(2.2)$$

for all $n \geqslant n_0^1$ and $p \geqslant p_0$. It is enough to prove that $\exists\ n'^{1}_{0}$ such that for $n > n'^{1}_{0}$, $0 \leqslant p \leqslant p_0$

$$\frac{1}{n} \sum_{j=p}^{p+n-1} | x_j - L | < \epsilon. \qquad \ldots(2.3)$$

Since, taking $n_0 = \max(n_0^1, n'^{1}_{0})$, (2.3) will hold for $n \geqslant n_0$ and for all p, which gives the result.

Once p_0 has been chosen, p_0 is fixed, so

$$\sum_{j=0}^{p_0-1} | x_j - L | = M \text{ (say)}. \qquad \ldots(2.4)$$

Now, taking $0 \leqslant p \leqslant p_0$ and $n > p_0$, we have

$$\frac{1}{n}\sum_{j=p}^{p+n-1} | x_j - L | = \frac{1}{n}\left(\sum_{j=p}^{p_0-1} + \sum_{j=p_0}^{p+n-1}\right) | x_j - L | \qquad ...(2.5)$$

$$\leqslant \frac{M}{n} + \frac{1}{n}\sum_{j=p_0}^{p_0+n-1} | x_j - L | \text{ (from 2.4)}$$

$$\leqslant \frac{M}{n} + \frac{\epsilon}{2} \text{ (by 2.2)}$$

Taking, n sufficiently large, we can made

$$\frac{M}{n} + \frac{\epsilon}{2} < \epsilon \qquad ...(2.6)$$

which gives (2.3) and hence the result.

PROOF OF THEOREM 1: Let $\{x_j\} \in | AC_\theta |$

$\Rightarrow$ Given $\epsilon > 0$, $\exists\ r_0$, and L such that

$$\frac{1}{h_r}\sum_{j=q}^{q+h_r-1} | x_j - L | < \epsilon \qquad ...(2.7)$$

for $r > r_0$, and $q = K_{r-1} + 1 + i,\ i \geqslant 0$.

Let $n \geqslant h_r$, write

$$n = mh_r + \theta \qquad ...(2.8)$$

where, $0 \leqslant \theta \leqslant h_r$, m is an integer. Since $h \geqslant h_r$, $m \geqslant 1$. Now

$$\frac{1}{n}\sum_{j=q}^{q+n-1} | x_j - L | \leqslant \frac{1}{n}\sum_{j=q}^{q+(m+1)h_r-1} | x_j - L |$$

$$= \frac{1}{n}\sum_{u=0}^{m}\sum_{j=q+uh_r}^{q+(u+1)h_r-1} | x_j - L |$$

$$\leqslant \frac{m+1}{n} h_r \epsilon \qquad ...(2.9)$$

$$\leqslant \frac{2mh_r\epsilon}{n} \quad (m \geqslant 1).$$

For $\frac{h_r}{n} \leqslant 1$, since, $\frac{mh_r}{n} \leqslant 1$

$$\frac{1}{n} \sum_{j=q}^{q+n-1} | x_j - L | \leqslant 2\epsilon.$$

Then by Lemma 1, $| AC_\theta | \Rightarrow | AC |$. It is trivial that $| AC | \Rightarrow |AC_\theta|$ for every θ. Hence, we have the result.

Theorem 2— (a) For some, θ, $AC_\theta =\!/\!\Rightarrow l\,\infty$;

(b) For every θ, $AC_\theta \cap l\,\infty \Leftrightarrow AC$.

In order to prove this theorem, we require the following lemma.

Lemma 2—Suppose for given $\epsilon > 0$, $\exists\, n_0, p_0$ such that

$$\frac{1}{n} \Big| \sum_{j=p}^{p+n-1} x_j - L \Big| < \epsilon$$

for all $n \geqslant n_0$, $p \geqslant p_0$. Then $\{x_j\} \in AC$.

PROOF : Let $\epsilon > 0$ be given. Choose n_0', p_0 such that

$$\frac{1}{n} \Big| \sum_{j=p}^{p+n-1} x_j - L \Big| < \epsilon/2 \qquad \text{...(2.10)}$$

for $n \geqslant n_0'$ $p \geqslant p_0$.

As in Lemma 1, it is enough to show, $\exists\, n_0''$ such that for $n \geqslant n_0''$, $0 \leqslant p \leqslant p_0$, we have

$$\frac{1}{n} \Big| \sum_{j=p}^{p+n-1} x_j - L \Big| < \epsilon. \qquad \text{...(2.11)}$$

Since p_0 is fixed. Let

$$\sum_{j=0}^{p_0-1} | x_j - L | = M \text{ (say)}. \qquad \text{...(2.12)}$$

Now, let $0 \leqslant p \leqslant p_0$, and $n > p_0$, then

$$\frac{1}{n} \Big| \sum_{j=p}^{p+n-1} x_j - L \Big| \leqslant \frac{1}{n} \sum_{j=p}^{p_0-1} | x_j - L | + \frac{1}{n} \Big| \sum_{j=p_0}^{p+n-1} x_j - L \Big|$$

(*equation continued on p.* 69)

$$\leqslant \frac{M}{n} + \frac{1}{n} \left| \sum_{j=p_0}^{p_0+n+p-p_0-1} x_j - L \right|, \qquad ...(2.13)$$

Let $n - p_0 > n'_0$ Then for $0 \leqslant p < p_0$, we have $n + p - p_0 \geqslant n'_0$. From (2.10)

$$\frac{1}{n+p+p_0} \left| \sum_{p=p_0}^{p_0+n+p-p_0} x_j - L \right| < \epsilon/2 \qquad ...(2.14)$$

From (2.13) and (2.14)

$$\frac{1}{n} \left| \sum_{j=p}^{p+n-1} x_j - L \right| \leqslant \frac{M}{n} + \frac{n+p-p_0}{n} \epsilon/2$$

$$\leqslant \frac{M}{n} + \epsilon/2$$

$< \epsilon$, for sufficiently large n.

Hence the result,

Proof of Theorem 2 : (a) It is enough to show $AC_\theta \not\Rightarrow l_\infty$ when, K_r is even for all r. Let

$$x_k = (-1)^k K^\lambda. \qquad ...(2.15)$$

Where λ is a constant, with $0 < \lambda < 1$. Then

$$\sum_{j=q}^{q+h_r-1} x_j,\ q \geqslant 0 \qquad ...(2.16)$$

will contains an even number of terms. So we can bracket the terms of the sum in pairs of consecutive terms without there being a term left over. Since, the sum of two consecutive terms is $O(K^{\lambda-1})$, it is straight forward matter to verify that $\{x_j\} \in AC_\theta$ with $L = 0$. But $\{x_j\}$ is not bounded.

Remark : $AC_\theta \Rightarrow l\,\infty$, whenever the set of values for h_r includes all positive integers in any order.

(b) Let $\{x_j\} \in AC_\theta \cap l\,\infty$. For $\epsilon > 0$, $\exists\ r_0$ and q_0 such that

$$\frac{1}{h_r} \left| \sum_{j=q}^{q+h_r-1} x_j - L \right| < \epsilon/2 \qquad ...(2.17)$$

for $r \geqslant r_0$, $q \geqslant q_0$. $q = K_{r-1} + 1 + i$, $i \geqslant 0$,

Now, let $n \geqslant h_r$, m is an integer greater than equal to 1. Then

$$\frac{1}{n} \left| \sum_{j=q}^{q+n-1} x_j - L \right| \leqslant \frac{1}{n} \sum_{\mu=0}^{m-1} \left| \sum_{j=q+\mu h_r}^{q+(\mu+1)h_r-1} x_j - L \right|$$

$$+ \frac{1}{n} = \sum_{j=q+mh_r}^{q+n-1} | x_j - L | \qquad \ldots(2.18)$$

Since $\{x_j\} \in l\,\infty$, let all j, $| x_j - L | < M$ (say).

So, from (2.17) and (2.18)

$$\frac{1}{n} \left| \sum_{j=q}^{q+n-1} x_j - L \right| \leqslant \frac{1}{n} m.\, h_r\, \epsilon/2 + \frac{Mh_r}{n}. \qquad \ldots(2.19)$$

For, $\frac{h_r}{n} \leqslant 1$, since $\frac{mh_r}{n} \leqslant 1$, and $\frac{Mh_r}{n}$ can be made less than $\epsilon/2$ by taking n sufficiently large.

So,

$$\frac{1}{n} \left| \sum_{j=q}^{q+n-1} x_j - L \right| < \epsilon \text{ for } r \geqslant r_0,\, q > q_0.$$

Hence, by Lemma 2, $AC_\theta \cap l\,\infty \Rightarrow AC$.

It is trivial that, $AC \Rightarrow AC_\theta \cap l\,\infty$.

§3. In view of the ideas of uniformly and well distributed modulo 1 of Weyl[9] and Petersen[8] and result in §2, it is natural to define uniformly and well distributed modulo 1, of the sequence of real numbers, over the lacunary sequence θ as follows :

Let $x = (x_K)$ be a sequence of real numbers such that

$$0 \leqslant x_k < 1, \forall\, k \geqslant 0. \qquad \ldots(3.0)$$

$0 \leqslant a < b \leqslant 1$ and $I_{(a,b)}$ be the characteristic function of interval $[a,b)$. We record the definitions of Weyl[9] and Petersen[8]:

Definition A[9]—A sequence 'x' is called uniformly distributed (abbreviated U.d.) if for every pair of a, b.

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} I_{[a,b)}(x_K) = b - a. \qquad \ldots(3.1)$$

Definition B[8]—A sequence 'x' is called well distributed (W.d.) if for each pairs of a, b

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=i}^{n+i-1} I_{[a,b)} (x_K) = b - a \qquad \ldots(3.2)$$

uniformly in $i \geqslant 0$.

Now, we define the well and uniform distributed over the lacunary sequence θ as follows :

Definition 1—A sequence 'x' is called uniformly distributed over θ ($u.\ d_\theta.$) if for every pairs of a, b

$$\lim \frac{1}{h_r} \sum_{K\in I_r} I_{[a,b)} (x_K) = b - a. \qquad \ldots(3.3)$$

Definition 2—A sequence x is called well distributed over θ (W.d$_\theta$.) if, for every pair of a,b

$$\lim_{r\to\infty} \frac{1}{h_r} \sum_{K\in I_r} I_{[a,b)} (x_{k+i}) = b - a \qquad \ldots(3.4)$$

uniformly in $i \geqslant 0$.

Let $A = (a_{nk}(j))$, $n, k, j \geqslant 0$ be the generalised three parametric real matrix with the following conditions.

$$\|A\| = \sup_{n,j} \sum_{k=0}^{\infty} | a_{nk}(j) | < \infty \qquad \ldots(3.5)$$

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{nk}(j) = 1, \text{ uniformly in } j. \qquad \ldots(3.6)$$

A is almost positive

i. e.

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{nk}^{-(j)} = 0 \text{ uniformly in } j \qquad \ldots(3.7)$$

where

$$a_{nk}^{-}(j) = \max(-a_{n,k}(j), 0).$$

$$a^{+}_{n,k} = \max (a_{n,k} (j)\ 0).$$

Definition C[7]—A sequence is said to have A-uniform asympotic distribution function $g\,[A$ – u.a.d.f., $g)$, if there exists, $g : [0, 1] \to \mathbb{R}$ such that

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k} (j)\, I_{[0,b)} (x_K) = g\,(b). \qquad \ldots(3.8)$$

Uniformly in $j \geqslant 0$.

Note that, taking $g\,(t) = t,\ 0 \leqslant t \leqslant 1$ and

for all $j \geqslant 0$

$$a_{n,k} (j) = \begin{cases} \frac{1}{n},\ 0 \leqslant k \leqslant n-1, \\ 0 \text{ elsewhere} \end{cases} \qquad \ldots(3.9)$$

$$a_{n,k} (j) = \begin{cases} \frac{1}{n},\ j \leqslant K \leqslant j+n-1 \\ 0 \text{ elsewhere} \end{cases} \qquad \ldots(3.10)$$

$$a_{n,k} (j) = \begin{cases} \frac{1}{h_r},\ K \in I_r,\ r = n \\ 0 \text{ elsewhere} \end{cases} \qquad \ldots(3.11)$$

$$a_{n,k} (j = \begin{cases} \frac{1}{h_r},\ K_{r-1} + j + 1 \leqslant K \leqslant K_r + j,\ r = n \\ 0 \text{ elsewhere.} \end{cases} \qquad \ldots(3.12)$$

Then definitions A, B,1,2 are the particular case of Definition C. Now we record the following from Das and Patel[7].

Theorem B—Let g be an non-negative, non decreasing bounded continuous function from [0,1) into $\mathbb{R}$. Then a sequence (x_k) is A – *u.a.d.f.g*, if and only if, for every real valued continuous function f.

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k} (j)\, f\,(x_k) = \int_0^1 f\,(t)\, dg\,(t) \qquad \ldots(3.13)$$

uniformly in $j \geqslant 0$.

Theorem C—Let g be a non-negative, non decreasing bounded continuous function from [0, 1] into $\mathbb{R}$. Then a sequence (x_k) is A – U.a.d,f,g, if and only if

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k} (j)\ e^{2\pi i h x_k} = \int_0^1 e^{2\pi i h x}\, dg\,(x). \qquad \ldots(3.14)$$

Uniformly in $j \geqslant 0$, for all non zero integer h, and $i = \sqrt{-1}$.

Similar to Weyl[9] and Petersen[8] and from B and C, We have the followings :

Theorem 3— (a) 'x' is $u.\ d_\theta$. if and only if

$$\lim_{r\to\infty} \frac{1}{h_r} \sum_{K\in I_r} f(x_k) = \int_0^1 f(x).\ dx \qquad \text{.. (3.15)}$$

for every Riemann integrable function f on [0, 1].

(b) 'x' is $W.\ d_\theta$. if and only if

$$\lim_{r\to\infty} \frac{1}{h_r} \sum_{K\in I_r} f(x_{k+l}) = \int_0^1 (x).\ dx \qquad \text{...(3.16)}$$

uniformly in $i \geqslant 0$, for every Riemann integrable function f on [0, 1],

(c) 'X' is $u.\ d_\theta$, if and only if

$$\lim_{r\to\infty} \frac{1}{h_r} \sum_{K\in I_f} e^{2\pi\, ihx_k} = 0 \qquad \text{...(3.17)}$$

for all integers $h \neq \sqrt{-1}$

(d) 'x' is $W.\ d_\theta$. if and only if

$$\lim_{r\to\infty} \frac{1}{h_r} \sum_{|K\in I_r} e^{2\pi ihx_{k+j}} = 0 \qquad \text{...(3.18)}$$

uniformly in $j \geqslant 0$, for all integers $h \neq 0$, $i = \sqrt{-1}$.

Theorem 4— (a) X is $u.\ d. \Rightarrow X$ is $u.\ d\theta$, if

$$\liminf_{r} q_r > 1 \qquad \text{...(3.19)}$$

(b) X is $U, d_\theta \Rightarrow X$ is $U.\ d.$, if

$$\lim_{r} q_r = 1 \qquad \text{...(3.20)}$$

(c) For every θ, (i) $W.d. \Rightarrow U.\ d_\theta$; (ii) $W.\ d_\theta \Leftrightarrow W.\ d.$

Proof (a) : Follows from Theorem A (d). Since $I_{[a,b]}(x)$ takes the value 0 or 1 only, we have

$$\min[(b-a),\ 1-(b-a)] \leqslant |\, I_{[a,b]}(x) - (b-a)\, |$$
$$\leqslant \max[(b-c),\ 1-(b-a)] \qquad \text{...(3.21)}$$

So $\{I_{[a\,b)}(x) - (b - a)\} \in l_\infty$. Now ($b$) follows from Theorem A (f) (c (i)) follows from Theorem A (g) and (c (ii)) follows from Theorem 2 (b).

Acknowledgement

The authors are grateful to the refree for his careful reading of the paper and valuable comments.

References

1. G. Das and S. Mishra, Sublinear functional and a class of conservative matrices. (Under communication).
2. A. R. Freedman, J. J. Sember and M. Raphael, *Proc. Lond. Math. Soc.* 37 (1978), 508–20.
3. G. G. Lorentz, *Acta Math.* 80 (1948), 167–90.
4. I. J. Maddox, *Proc. Camb. Phil. Soc.* 64 (1958) 335–40.
5. I. J. Maddox, *Proc. Camb. Phil. Soc.* 83 1978.
6. I. J. Maddox, *Math. Proc. Camb. Phil. Soc.* 85 (1979), 345–50.
7. B. K. Patel, Ph. D. thesis submitted to Sambalpur University 1982.
8. G. M. Petersen, *Quart. J. Math* (2) 7 (1956), 188–91.
9. H. Weyl, *Nachr. Ges. Wiss. Gottingen, Math. Phys.* 11 (1914), 234–44.

Indian J. pure appl. Math., 20 (1) : 75-88, January 1989

NOTE ON THE SMALL VIBRATION OF BEAMS WITH VARYING YOUNG'S MODULUS CARRYING A CONCENTRATED MASS DISTRIBUTION

P. K. CHAUDHURI

Department of Applied Mathematics, University College of Science
92, *Acharya Prafulla Chandra Road, Calcutta* 700 009

AND

SUBRATA DATTA

Behala Parnasree Bidyamandir, Calcutta 700 060

(*Received* 8 *April* 1987; *after revision* 4 *April* 1988)

The problem of vibration of beams or rods with varying Young's modulus has been solved assuming that the beam or rod carries a concentrated mass distribution. The effect of concentrated mass has been described by introducing Dirac δ-function in the differential equation. The methods of separation of variables and the Laplace transform have been used. Forced vibration problem has been solved after deriving the orthogonality relation for a variable density beam.

1. INTRODUCTION

The problem of vibration of beams or rods is of great importance in the theory of elasticity. The most convenient technique, to solve the homogeneous partial differential equation describing the free vibration of beams or rods, is to use the method of separation of variables. This yields a pair of ordinary differential equations. The solution of the differential equation containing spatial variables, together with the boundary conditions, gives a set of eigen functions. On the other hand, the eigen function expansion method usually gives a solution of the general initial value problem either of free vibration or of forced vibration.

Timoshenko[1] solved the problem of longitudinal vibration of a rod in which a mass was attached to the end. Hoppmann[2] studied the transverse vibration of a beam carrying a mass at the middle. His discussion did not include the general initial value problem. Moreover, neither Timoshenko[1] nor Hoppmann[2] used eigen function method, perhaps due to the difficulty in deriving the orthogonality relation. Morgan[3] showed an easy method of obtaining orthogonality relation for the cases where a Dirac δ-function described the effect of concentrated mass.

Introducing the method discussed by Morgan[3], Chen[4] obtained the solution of the vibration of beams or rods carrying a concentrated mass. Pan[5] investigated the

transverse vibration of a beam carrying a system of heavy bodies. Tseitlin[6] considered the longitudinal oscillations of a semi infinite rod with a mass fixed at the end. Howson and Williams[9] derived a solution in the matrix form for the natural frequencies of vibration of a structural frame. Chang and Juan[12] derived a set of equations for the free vibration of an inclined bar with an end constraint. Chen[13] presented the general dynamic stiffness matrix of a Timoshenko beam for transverse vibrations. Park[15] studied the dynamic stability of a uniform free-free Timoshenko beam under the action of controlled force.

In all the problems mentioned above the beam or the rod was considered to be homogeneous in the sense that Young's modulus of the beam or the rod was constant throughout. But as there are plenty of materials in nature which are not elastically homogeneous, the problems of vibrations of beams or rods with varying Young's modulus or with varying density have also received the attention of numerous investigators. Marangoni *et al.*[7] obtained the transverse vibrational frequency of uniform beams in which a steady state temperature distribution introduced a coordinate dependent elastic modulus. Nayfeh[8] derived the eigen values and corresponding eigen functions of free longitudinal vibration of a finite rod whose Young's modulus and mass density were assumed to vary continuously along the rod. Among the many other problems of vibrations of elastically nonhomogeneous bodies mention may be made of the interesting works of Shahinpoor[10], Farshad and Ahmadi[11] and Ercoli and Laura[14].

The purpose of the present investigation is to apply the techniques of Chen[4] and Pan[5] to a class of beams or rods which are elastically nonhomogeneous in the sense that the Young's modulus of the beams or rods considered here, is a function of position.

In our investigation of the problem we have first considered the simple case of vibration where only a single mass is placed at the middle of the rod and then the more general case where a system of heavy bodies have been placed at different points on the rod and driving forces and driving moments have also been taken into account. We refer these two cases as Case I and Case II respectively.

Thus the present problem may be considered as an extension of the works of Chen[4] and also of Pan[5].

2. Formulation and the Solution of the Problem in Case I

Let us consider the free vibration of a simply supported beam of nonhomogeneous material having length $2L$. Let the beam carry a concentrated mass M at its middle. The equation of motion, for Euler Bernoulli's description of the beam, may be written as,

$$\frac{\partial^2}{\partial x^2}\left(EI \frac{\partial^2 y}{\partial x^2} \right) + \left[\rho + M\,\delta\,(x - L) \right] \frac{\partial^2 y}{\partial t^2} = 0 \qquad \ldots(1)$$

where E is the Young's modulus and ρ is the mass per unit length of the beam, and the δ-function satisfies the relation,

$$\int_0^{2L} \delta(x - L)\, dx = 1. \qquad \ldots(2)$$

The boundary conditions of the problem are,

$$y(0, t) = y(2L, t) = y''(0, t) = y''(2L, t) = 0 \qquad \ldots(3)$$

where the primes indicate differentiation of y with respect to x.

Thus our problem is to find y from (1) subject to the boundary conditions (3).

If we assume a series solution of (1) in the form,

$$y = \sum_{n=1}^{\infty} c_n^* \psi_n(x) \sin p_n t \qquad \ldots(4)$$

we get the equation for $\psi_n(x)$ as

$$\frac{d^2}{dx^2}\left(EI \frac{d^2 \psi_n}{dx^2} \right) - p_n^2 \left[\rho + M \delta(x - L) \right] \psi_n = 0 \qquad \ldots(5)$$

subject to the boundary conditions

$$\psi_n(0) = \psi_n(2L) = \psi_n''(0) = \psi_n''(2L) = 0. \qquad \ldots(6)$$

Let the Young's modulus E and the density ρ of the beam be functions of position given by

$$E = E_0 e^{\alpha x} \qquad \ldots(7)$$

$$\rho = \rho_0 e^{\alpha x}$$

α being any real quantity having the dimension of reciprocal of length.

Equations (5) and (6) define an eigen value problem. Let us proceed to solve it by the method of Laplace transform. Denoting the transformed quantity by a bar over the corresponding function, we may write,

$$\bar{\psi}_n = \frac{(s + \alpha)^2}{s^4 + 2s^3 \alpha + s^2 \alpha^2 - k_n^4} \psi_n'(0) + \frac{1}{s^4 + 2s^3 \alpha + s^2 \alpha^2 - k_n^4} \psi_n'''(0)$$

$$+ \frac{M p_n^2 e^{-L\alpha} \psi_n(L)}{E_0 I} \frac{e^{-Ls}}{s^4 + 2s^3 \alpha + s^2 \alpha^2 - k_n^4} \qquad \ldots(8)$$

where $k_n^4 = \rho_0 \, p_n^2 / E_0 I$ and s is the Laplace transform parameter.

While deriving equation (8) two of the boundary conditions (6) namely $\psi_n(0)=0$ and $\psi_n''(0) = 0$ have been used.

Taking inversion of (8) we get,

$$\psi_n(x) = \psi_n'(0) f(x) + \psi_n'''(0) g(x) + \frac{M p_n^2 e^{-L\alpha} \psi_n(L)}{E_0 I} h(x) \qquad \ldots(9)$$

where

$$f(x) = \exp(-\alpha x/2) [A_1 \cos \beta_n x + C_1 \cosh \gamma_n x + \xi_1 \sin \beta_n x$$
$$+ \eta_1 \sinh \gamma_n x]$$

$$g(x) = \exp(-\alpha x/2) [A_2 \cos \beta_n x + C_2 \cosh \gamma_n x + \xi_2 \sin \beta_n x$$
$$+ \eta_2 \sinh \gamma_n x]$$

$$h(x) = g(x - L) H(X - L)$$

$$\xi_i = \left(B_i - \frac{A_i \alpha}{2}\right)\frac{1}{\beta_n}, \eta_i = \left(D_i - \frac{C_i \alpha}{2}\right)\frac{1}{\gamma_n} \quad (i = 1, 2)$$

$$A_1 = -\frac{\alpha}{2k_n^2}, C_1 = \frac{\alpha}{2k_n^2}$$

$$B_1 = \tfrac{1}{2}\left(1 - \frac{\alpha^2}{k_n^2}\right), D_1 = \tfrac{1}{2}\left(1 + \frac{\alpha^2}{k_n^2}\right)$$

$$A_2 = C_2 = 0$$

$$B_2 = -\frac{1}{2k_n^2}, D_2 = \frac{1}{2k_n^2}$$

$$\beta_n^2 = k_n^2 - \frac{\alpha^2}{4}, \gamma_n^2 = k_n^2 + \frac{\alpha^2}{4}.$$

$H(x - L)$ being a unit step function at $x = L$.

The constants $\psi_n'(0)$ and $\psi_n'''(0)$ in (9) may be determined from the remaining two boundary conditions of (6) i.e., $\psi_n(2L) = 0$ and $\psi_n''(2L) = 0$. The results obtained are

$$\psi_n'(0) = L_1 \frac{Mp_n^2 e^{-L\alpha} \psi_n(L)}{E_0 I} \qquad \ldots(10)$$

$$\psi_n'''(0) = M_1 \frac{Mp_n^2 e^{-L\alpha} \psi_n(L)}{E_0 I}$$

where

$$L_1 = \frac{g(2L)\, h''(2L) - g''(2L)\, h(2L)}{f(2L)\, g''(2L) - f''(2L)\, g(2L)}$$

$$M_1 = \frac{f''(2L)\, h(2L) - f(2L)\, h''(2L)}{f(2L)\, g''(2L) - f''(2L)\, g(2L)}.$$

Thus in view of (10) eqn. (9) becomes,

$$\psi_n(x) = \frac{Mp_n^2 e^{-L\alpha} \psi_n(L)}{E_0 I}\Big[L_1 f(x) + M_1 g(x) + h(x)\Big]. \qquad \ldots(11)$$

If we let $x = L$ in (11) the following equation results,

$$\psi_n(L)\Big[\; Mp_n^2\; e^{-L\alpha}\Big\{L_1 f(L) + M_1 g(L) + h(L)\Big\} - E_0\, I\Big] = 0. \ldots(12)$$

Since $\psi_n(L) = 0$ corresponds to the trivial solution of y, it follows that for nontrivial solutions

$$Mp_n^2\, e^{-L\alpha}\Big\{L_1 f(L) + M_1\, g(L) + h(L)\Big\} - E_0\, I = 0. \qquad \ldots(13)$$

Equation (13) gives an infinite set of eigenvalues p_n, each of which corresponds to the eigen function $\psi_n(x)$. Substituting their values in (4) we get the solution of the problem in this case. The arbitrary constants c_n^* are to be determined from the initial conditions of the problem.

At this stage it may be important to note the reduction of the function $\psi_n(x)$ in (9) and the equation (13) giving the eigen values p_n, as $\alpha \to 0$. Thus making $\alpha \to 0$, (9) and (13) reduce to

$$\psi_n(x) = \frac{\psi_n'(0)}{2k_n}\Big[\sin k_n x + \sinh k_n x\Big] + \frac{\psi_n'''(0)}{2k_n^3}\Big[\sinh k_n x - \sin k_n x\Big]$$

$$+ \frac{Mp_n^2\, \psi_n(L)}{2E_0\, Ik_n^3} H(x - L)\Big[\sinh k_n(x - L) - \sin k_n(x - L)\Big]$$

$$\ldots(9a)$$

and

$$Mp_n^2 (\tan k_n L - \tanh k_n L) - 4 E_0 I k_n^3 = 0 \qquad \text{...(13a)}$$

which are in complete agreement with the equations (8) and (12) respectively, obtained by Chen[4] for an elastically homogeneous beam.

2.1. *The Orthogonality Relation*

It has been suggested by Morgan[3] that if the function $\rho + M\delta (x - L)$ be treated as a weighting function then the eigen functions $\psi_n (x)$ are orthogonal with respect to this weighting function.

Let us assume

$$\epsilon (x) = \rho + M\delta (x - L) \qquad \text{...(14)}$$

then

$$\int_0^{2L} \epsilon (x) \psi_n (x) \psi_m (x)\, dx = \frac{E_0 I}{p_n^2 - p_m^2} \int_0^{2L} e^{\alpha x} \Big[\psi_m (x)\, \psi_n'''' (x) - \psi_n (x) \psi_m'''' (x)$$

$$+ 2\alpha\, \psi_m (x) \psi_n''' (x) - 2\alpha\, \psi_n (x)\, \psi_m''' (x)$$

$$+ \alpha^2 \psi_m (x)\, \psi_n'' (x) - \alpha^2 \psi_n (x)\, \psi_m'' (x) \Big] dx = 0$$

$$\text{for } m \neq n. \qquad \text{...(15)}$$

Verification of eqn. (15) is a very simple task.

Now if we substitute (14) in (15) we get,

$$\int_0^{2L} e^{\alpha x} \psi_n (x) \psi_m (x)\, dx = -\frac{M}{\rho_0} \psi_n (L) \psi_m (L). \qquad \text{...(16)}$$

Let

$$\int_0^{2L} \epsilon (x) \psi_n^2 (x)\, dx = \frac{E_0 I}{p_n^2} \int_0^{2L} e^{\alpha x} \Big[\psi'' (x) \Big]^2 dx = N_n. \qquad \text{...(17)}$$

N_n is the normalization constant for the system of eigen functions.

2.2. *The Forced Vibration*

Let us determine the response of the beam to an impulsive load.

For a unit impulsive load the differential equation of the forced vibration may be written as,

$$\frac{\partial^2}{\partial x^2}\left(EI\, \frac{\partial^2 y}{\partial x^2} \right) + \Big[\rho + M\, \delta\, (x - L) \Big] \frac{\partial^2 y}{\partial t^2} = \delta\, (t)\, \delta\, (x - L). \qquad \text{...(18)}$$

Let

$$y = \sum_{n=1}^{\infty} q_n(t)\, \psi_n(x). \qquad \text{...(19)}$$

$\psi_n(x)$ is the eigen function which we have already determined and the function $q_n(t)$ is to be determined. Substituting (19) in (18), multiplying the resulting equation by ψ_m and then integrating from 0 to $2L$ we obtain,

$$\int_0^{2L} \psi_m \frac{\partial^2}{\partial x^2} \left(EI \sum_{n=1}^{\infty} \psi_n'' q_n\right) dx + \int_0^{2L} \psi_m \epsilon(x) \left(\sum_{n=1}^{\infty} \psi_n \ddot{q}_n\right) dx$$

$$= \delta(t) \int_0^{2L} \psi_m\, \delta(x - L)\, dx. \qquad \text{...(20)}$$

Now ψ_n satisfies eqn. (5) as well as the orthogonality relation. Hence eqn. (20) reduces to

$$\ddot{q}_n + p_n^2 q_n = \frac{\psi_n(L)}{N_n} \delta(t) \qquad n = 1, 2, \ldots \qquad \text{...(21)}$$

The particular solution of equation (21) is given by

$$q_n = \frac{\psi_n(L)}{N_n p_n} \sin p_n t. \qquad \text{...(22)}$$

For a general forcing function $F(t)$ we get, by convolution, the particular solution as,

$$y(t) = \sum_{n=1}^{\infty} \frac{\psi_n(L)\, \psi_n(x)}{N_n p_n} \int_0^t \sin p_n(t - \xi)\, F(\xi)\, d\xi. \qquad \text{...(23)}$$

Here it should be noted that by virtue of the orthogonality relation the solution of the initial-value problem can easily be obtained by expanding the initial conditions into series of eigen functions.

3. Formulation and the Solution of the Problem in Case II

When a uniform beam, carrying a system of heavy bodies placed at the points $x = a_l$, is subjected to a system of driving forces $F_j(t)$ and driving moments $G_k(t)$ acting at points b_j and c_k respectively, its equation of motion may be written as,

$$\frac{\partial^2}{\partial x^2}\left(EI \frac{\partial^2 y}{\partial x^2} \right) - \frac{\partial}{\partial x}\left[\sum_{l=1}^{R} J_l\, \delta(x - a_l) \frac{\partial^3 y}{\partial x\, \partial t^2} \right]$$

$$+ \left[\rho + \sum_{l=1}^{R} M_l\, \delta(x - a_l) \right] \frac{\partial^2 y}{\partial t^2} = \sum_{j=1}^{S} F_j(t)\, \delta(x - b_j)$$

(*equation continued on p. 82*)

$$+ \sum_{k=1}^{T} G_k (t)\, \delta' (x - c_k). \qquad \ldots(24)$$

In eqn. (24) M_i and J_i respectively denote the mass and the rotary inertia of the ith body; R, S, T are the total number of heavy bodies, the total number of driving forces and the total number of driving moments respectively.

For free vibration of the beam the right-hand side of (24) is zero. In this case if we assume a series solution in the form

$$y (x, t) = \sum_{n=1}^{\infty} \psi_n (x) \Big[P_n \sin p_n t + Q_n \cos p_n t \Big] \qquad \ldots(25)$$

then the reduced homogeneous equation, when (7) is taken into account, yields the following spatial equation,

$$\frac{\partial^2}{\partial x^2} \Big[E_0\, I\, e^{\alpha x} \psi_n'' \Big] + p_n^2 \left\{ \Big[\sum_{i=1}^{R} J_i\, \delta(x - a_i)\, \psi_n' \Big]' \right.$$

$$\left. - \Big[\rho_0\, e^{\alpha x} + \sum_{i=1}^{R} M_i\, \delta(x - a_i) \Big] \psi_n \right\} = 0. \qquad \ldots(26)$$

Thus in this case our problem is to find ψ_n from (26) subject to the boundary conditions (6).

As in Case I, noting the boundary conditions (6) and taking Laplace transform of (26) we obtain,

$$\bar{\psi}_n = \frac{(s + \alpha)^2}{s^4 + 2s^3 \alpha + s^2 \alpha^2 - k_n^4} \psi_n' (0) + \frac{1}{s^4 + 2s^3 \alpha + s^2 \alpha^2 - k_n^4} \psi_n'''(0)$$

$$- \frac{s}{s^4 + 2s^3 \alpha + s^2 \alpha^2 - k_n^4} \frac{p_n^2}{E_0 I} \sum_{i=1}^{R} J_i\, \psi_n' (a_i)\, e^{-a_i \alpha}.\, e^{-a_i s}$$

$$+ \frac{1}{s^4 + 2s^3 \alpha + s^2 \alpha^2 - k_n^4} \frac{p_n^2}{E_0 I} \sum_{i=1}^{R} M_i\, \psi_n (a_i)\, e^{-a_i \alpha}.\, e^{-a_i s}.$$

$$\ldots(27)$$

Taking inversion of (27) we get,

$$\psi_n (x) = \psi_n' (0) f (x) + \psi_n''' (0) g (x) - \frac{p_n^2}{E_0 I} \sum_{i=1}^{R} J_i\, \psi_n' (a_i)\, e^{-a_i \alpha}.\, f_1 (x - a_i)$$

(*equation contiued on p.* 83)

$$\times \frac{p_n^2}{E_0 I} \sum_{i=1}^{R} M_i \psi_n (a_i) e^{-a_i \alpha} g_1 (x - a_i) \qquad \ldots(28)$$

where

$$f_1(x) = e^{-\frac{\alpha}{2}x} H(x) [A_3 \cos \beta_n x + C_3 \cosh \gamma_n x$$
$$+ \xi_3 \sin \beta_n x + \eta_3 \sinh \gamma_n x]$$

$$g_1(x) = g(x) H(x)$$

$$\xi_3 = \left(B_3 - \frac{A_3 \alpha}{2}\right) \frac{1}{\beta_n} \qquad \eta_3 = \left(D_3 - \frac{C_3 \alpha}{2}\right) \frac{1}{\gamma_n}$$

$$A_3 = -\frac{1}{2k_n^2}, \; C_3 = \frac{1}{2k_n^2}$$

$$B_3 = D_3 = 0.$$

If we impose remaining two of the boundary conditions (6) namely $\psi_n\ (2L) = 0$ and $\psi_n''\ (2L) = 0$ on (28) we obtain the following equations :

$$\psi_n'\ (0)\ f(2L) + \psi_n'''\ (0)\ g(2L) - \frac{p_n^2}{E_0 I} \sum_{i=1}^{R} J_i \psi_n'\ (a_i)\ e^{-a_i \alpha} f_1 (2L - a_i)$$

$$+ \frac{p_n^2}{E_0 I} \sum_{i=1}^{R} M_i \psi_n (a_i) e^{-a_i \alpha} g_1 (2L - a_i) = 0 \qquad \ldots(29)$$

and

$$\psi_n'\ (0) f''\ (2L) + \psi_n'''\ (0)\ g''\ (2L) - \frac{p_n^2}{E_0 I} \sum_{i=1}^{R} J_i \psi_n'\ (a_i)\ e^{-a_i \alpha}$$

$$\times\ f_1''\ (2L - a_i) + \frac{p_n^2}{E_0 I} \sum_{i=1}^{R} M_i \psi_n (a_i)\ e^{-a_i \alpha}$$

$$\times\ g_1''\ (2L - a_i) = 0. \qquad \ldots(30)$$

Substituting $x = a_q - \epsilon$ in (28) and its derivative, and letting $\epsilon \to 0$, two systems of equations which represent the consistency conditions are obtained.

$$-\psi_n (a_q) + \psi_n'\ (0) f(a_q) + \psi_n'''\ (0)\ g(a_q) - \frac{p_n^2}{E_0 I} \sum_{i=1}^{q-1} J_i\ \psi_n'\ (a_i)$$

(equation continued on p. 84)

$$\times e^{-a_i\alpha} f_1(a_q - a_i) + \frac{p_n^2}{E_0 I} \sum_{i=1}^{q-1} M_i \psi_n(a_i) e^{-a_i\alpha} g_1(a_q - a_i) = 0,$$

$$q = 1, 2, 3, \ldots R \qquad \ldots(31)$$

$$-\psi_n'(a_q) + \psi_n'(0) f'(a_q) + \psi_n''' (0) g'(a_q) - \frac{p_n^2}{E_0 I} \sum_{i=1}^{q-1} J_i \psi_n'(a_i)$$

$$\times e^{-a_i\alpha} f_1'(a_q - a_i) + \frac{p_n^2}{E_0 I} \sum_{i=1}^{q-1} M_i \psi_n(a_i) e^{-a_i\alpha}$$

$$\times g_1'(a_q - a_i) = 0, q = 1, 2, 3, \ldots, R. \qquad \ldots(32)$$

Equations (29) to (32) form a system of homogeneous equations in $(2R + 2)$ unknowns $\psi_n'(0)$, $\psi_n'''(0)$, $\psi_n(a_i)$, $\psi_n'(a_i)$. For a nontrivial solution of the problem the determinant of the coefficient matrix X of the foregoing system of equations must vanish, leading to the frequency equation as

$$\det(X) = 0. \qquad \ldots(33)$$

Equation (33) is an equation determining k_n.

Substituting each k_n obtained from (33) back into equations (29) to (32) the values of $\psi_n'(0)$, $\psi_n'''(0)$, $\psi_n(a_i)$ and $\psi_n'(a_i)$ are determined. Corresponding to the set of eigen values k_n, the eigen functions are obtained from equation (28).

After obtaining $\psi_n(x)$, and p_n from k_n and substituting their values in (25) the solution of the free vibration is obtained. The constants P_n and Q_n are determined from the initial conditions $\psi(x, 0)$ and $\dot{\psi}(x, 0)$. This may be done by expanding these initial conditions in the obtained eigen functions and comparing the coefficients.

3.1. *Orthogonality condition and eigenfunction expansion*

We may write eqn. (26) in the following form :

$$-\frac{\partial^2}{\partial x^2}\left[e^{\alpha x}\psi_n''\right] = \lambda_n \left\{ \left[\sum_{i=1}^{R} J_i \delta(x - a_i) \psi_n' \right]' \right.$$

$$\left. - \left[\rho_0 e^{\alpha x} + \sum_{i=1}^{R} M_i \delta(x - a_i) \right] \psi_n \right\} \qquad \ldots(34)$$

where

$$\lambda_n = \frac{p_n^2}{E_0 I} = \frac{k_n^4}{\rho_0}.$$

For $m \neq n$ we may deduce from (34) the relation,

$$\lambda_n \int_0^{2L} \left\{ \left[- \sum_{i=1}^{R} J_i \, \delta(x - a_i) \, \psi_n' \right]' \psi_m \right.$$

$$\left. + \left[\rho_0 \, e^{\alpha x} + \sum_{i=1}^{R} M_i \, \delta(x - a_i) \right] \psi_n \, \psi_m \right\} dx$$

$$- \lambda_m \int_0^{2L} \left\{ \left[- \sum_{i=1}^{R} J_i \, \delta(x - a_i) \, \psi_m' \right]' \psi_n \right.$$

$$\left. + \left[\rho_0 \, e^{\alpha x} + \sum_{i=1}^{R} M_i \, \delta(x - a_i) \right] \psi_n \, \psi_m \right\} dx$$

$$= \int_0^{2L} e^{\alpha x} \Big[\psi_n'''' (x) \, \psi_m (x) - \psi_m'''' (x) \, \psi_n (x) + 2\alpha \, \psi_n''' (x) \, \psi_m (x)$$

$$2\alpha \, \psi_m''' (x) \, \psi_n (x) + \alpha^2 \, \psi_n'' (x) \, \psi_m (x)$$

$$- \alpha^2 \, \psi_m'' (x) \, \psi_n (x) \Big] dx = 0. \qquad \text{...(35)}$$

Since for $m \neq n$ the right-hand side of (35) is equal to zero and both the integrals on the left hand side yield same result, we may write eqn. (35) as

$$\int_0^{2L} \rho_0 \, e^{\alpha x} \, \psi_n \, \psi_m \, dx + \sum_{i=1}^{R} M_i \, \psi_n (a_i) \, \psi_m (a_i)$$

$$+ \sum_{i=1}^{R} J_i \, \psi_n' (a_i) \, \psi_m' (a_i) = 0. \qquad \text{...(36)}$$

In deriving (36) we have used the properties that,

$$\delta (x - a) = 0 \text{ for } x \neq a \qquad \text{...(37)}$$

and

$$\int_{-\infty}^{+\infty} \varphi (x) \, \delta (x - a) \, dx = \varphi (a).$$

For the special case, $J_i = 0$, $M_1 = M$ and $M_i = 0$ for $i \neq 1$

(36) reduces to,

$$\int_0^{2L} \rho_0 \, e^{\alpha x} \psi_n \psi_m \, dx + M \, \psi_n (a) \, \psi_m (a) = 0 \qquad \ldots(38)$$

which agrees with the condition (16) in Case I.

With the orthogonality condition (36) it can be shown that any arbitrary continuously differentiable function $h^* (x)$, defined in the interval $[0, 2L]$, has a Fourier-type expansion in the obtained eigen functions :

$$h^* (x) = \sum_{n=1}^{\infty} \zeta_n \, \psi_n (x)$$

with its Fourier coefficients ζ_n given by the formula,

$$\zeta_n = \frac{\int_0^{2L} \rho_0 \, e^{\alpha x} h^* \psi_n \, dx + \sum_{i=1}^{R} M_i \, h^* (a_i) \, \psi_n (a_i) + \sum_{i=1}^{R} J_i \, h^{*\prime} (a_i) \, \psi_n' (a_i)}{\int_0^{2L} \rho_0 \, e^{\alpha x} \psi_n^2 \, dx + \sum_{i=1}^{R} M_i \left[\psi_n (a_i) \right]^2 + \sum_{i=1}^{R} J_i \left[\psi_n' (a_i) \right]^2} \qquad \ldots(39)$$

3.2 *Forced vibration*

Let us expand $y (x, t)$ in the eigen functions,

$$y (x, t) = \sum_{n=1}^{\infty} q_n (t) \, \psi_n (x). \qquad \ldots(40)$$

After substituting (40) in (24) and then multiplying (24) by $\psi_m (x)$ and integrating we get,

$$E_0 \, I \sum_{n=1}^{\infty} q_n (t) \int_0^{2L} \left\{ \frac{\partial^2}{\partial x^2} \left[e^{\alpha x} \, \psi_n'' \right] \right\} \psi_m \, dx$$

$$+ \sum_{n=1}^{\infty} \ddot{q}_n (t) \int_0^{2L} \left\{ - \left[\sum_{i=1}^{R} J_i \, \delta (x - a_i) \, \psi_n' \right]' \right.$$

$$\left. + \left[\rho_0 \, e^{\alpha x} + \sum_{i=1}^{R} M_i \, \delta (x - a_i) \right] \psi_n \right\} \psi_m \, dx$$

(*equation continued on p.* 87)

$$= \sum_{j=1}^{S} F_j(t) \int_0^{2L} \delta(x - b_j)\, \psi_m \, dx$$

$$+ \sum_{k=1}^{T} G_k(t) \int_0^{2L} \delta'(x - c_k)\, \psi_m \, dx \qquad \ldots(41)$$

using (34) and the orthogonality condition (36) the general solution for the forced vibration is obtained as

$$y(x, t) = \sum_{n=1}^{\infty} (P_n \sin p_n t + Q_n \cos p_n t)\, \psi_n(x)$$

$$+ \sum_{n=1}^{\infty} \frac{\psi_n(x)}{p_n N_n} \int_0^{t} \Big[\sum_{j=1}^{S} \psi_n(b_j)\, F_j(\tau)$$

$$- \sum_{k=1}^{T} \psi_n'(c_k)\, G_k(\tau) \Big] \sin p_n (t - \tau)\, d\tau \qquad \ldots(42)$$

where

$$N_n = \int_0^{2L} \rho_0 e^{\alpha x} \psi_n^2 \, dx + \sum_{i=1}^{R} M_i \Big[\psi_n(a_i) \Big]^2$$

$$+ \sum_{i=1}^{R} J_i \Big[\psi_n(a_i) \Big]^2. \qquad \ldots(43)$$

If the forcing functions are sinusoidal $F_j(t) = F_j \sin \omega_j t$ and $G_k(t) = G_k \sin \omega_k t$, (42) may be integrated into,

$$y(x, t) = \sum_{n=1}^{\infty} (P_n \sin p_n t + Q_n \cos p_n t)\, \psi_n(x)$$

$$+ \sum_{n=1}^{\infty} \frac{\psi_n(x)}{p_n N_n} \Big\{ \sum_{j=1}^{S} \psi_n(b_j)\, F_j \Big(\frac{\omega_j}{\omega_j^2 - p_n^2} \sin p_n t$$

$$- \frac{p_n}{\omega_j^2 - p_n^2} \sin \omega_j t \Big) - \sum_{k=1}^{T} \psi_n'(c_k) G_k \Big(\frac{\omega_k}{\omega_k^2 - p_n^2} \sin p_n t$$

(equation continued on p. 88)

$$- \frac{p_n}{\omega_k^2 - p_n^2} \sin \omega_k t \Big) \Big\} . \qquad \ldots(44)$$

The constants P_n and Q_n, as mentioned earlier, are to be determined from the initial conditions.

Acknowledgement

The authors wish to thank referee for his useful comments and valuable suggestions for the improvement of the paper.

References

1. S. Timoshenko, *Vibration Problems in Engineering*. Third edition. McGraw-Hill Book Company, Inc., New York, 1954.
2. W. H. Hoppmann, *J Appl. Mech.* 19, *Trans. ASME*, 74 (1952), 301–307.
3. G. W. Morgan, *Quart. Appl. Math.* 11 (1953), 157–65.
4. Y. Chen, *J. Appl. Mech.* 30, *Trans. ASME*, 85 (1963), 310–11.
5. H. H. Pan, *J. Appl. Mech.* 32, *Trans. ASME*, 87 (1965), 434-36.
6. A. I. Tseitlin, *Stroit. Mekhan. i. Raschet. Sooruzh* No. 6, (1966), 40–45.
7. R. D. Marangoni, G. Fauconneau and L.A. Scipio, *Int. J. Eng. Sci.* 6 (1968), 637–59.
8. A. H. Nayfeh, *J. Appl. Mech., Trans. ASME Series* E 39 (1972), 595-97.
9. W. P. Howson and F. W. Williams, *J. Sound Vibr.* 26 (1973), 503–15.
10. M. Shahinpoor, *Iranian J. Sci. Tech.* 2 (1973), 257–78.
11. M. Farshad and G. Ahmadi, *Iranian J. Sci. Tech.* 3 (1974), 75–86.
12. C. H. Chang and Y. C. Juan, *J. Sound Vibr.* 101 (1985), 171–80.
13. Y. H. Chen, *Earth Quake Eng. Struct.* 15 (1987), 391–402.
14. L. Ercoli and P. A. A. Laura, *J. Sound Vibr.* 112 (1987), 447–54.
15. Y. P. Park, *J. Sound Vibr.* 113 (1987), 407–15.

Indian J. pure appl. Math., 20 (1) : 89-97, January 1989

FREE TORSIONAL VIBRATION OF A NONHOMOGENEOUS SEMIINFINITE SOLID CIRCULAR CYLINER

N. C. MONDAL

Patha-Bhavana, Visva-Bharati, Santiniketan 731235

(*Received* 9 *November* 1987)

Free torsional vibration of a nonhomogeneous semiinfinite solid circular cylinder with stress free and rigid boundary is solved in this paper. The nonhomogeneity of rigidity μ and density is chosen in the from

$$\mu = \mu_0 (1 + KZ)^n, \ \rho = \rho_0 (1 + KZ)^n.$$

The displacement and stress component for different values of n are numerically evaluated.

INTRODUCTION

Several attempts have been made to solve the problem of torsional oscillation in a circular cylinder by Love[1], Kolsky[2], Davies[3] and many others. Forced torsional vibration in a circular cylinder was studied e. g. by Mitra[4], Banerjee[5], Campbell and Tsao[6] and Mondal[7]. In all of these problems the curved boundary of the cylinder is taken to be stress-free. In many engineening applications, however, a cover or a shield is provided as a protection. Therefore the particles on the curved boundary are constrained from moving. Such a boundary condition may be mathematically accounted for by taking it as a rigid boundary. Recently free torsional vibration in an isotropic homogeneous infinitely long solid circular cylinder with rigid boundary was considered by Rao[8].

In the present paper an attempt has been made to extend the previous problem to the case of free torsional vibration of a semiinfinite solid circular cylinder with rigid boundary when the material of the cylinder is nonhomogeneous in rigidity and density. In order to compare we have also considered the analogous case when the curved surface is free from traction. The nonhomogeneity is taken in the form

$$\mu = \mu_0 (1 + KZ)^n, \ \rho = \rho_0 (1 + KZ)^n$$

where μ_0, ρ_0, K are constants, n is a rational number, z-is the axial co-ordinate. Solution is derived by the method of separation of variables. Explicit formulae are given for the stress component and particle displacement. These formulae are used to compute the particle displacement and stress component at a fixed depth and are compared graphically. As regards the case of rigid boundary nodal cylinder, the pattern is the same as in the case of stress free boundary.

FORMULATION AND SOLUTION OF THE PROBLEM

We consider a semiinfinite solid circular cylinder of radius 'a'. The origin is taken at the centre of one end of the cylinder, Z-axis is directed along the axis of the cylinder. We use cylindrical polar co-ordinates (r, θ, z) to specify the points of the cylinder. We first consider the case when the curved boundary surface is stress-free. It is assumed that the stress is distributed with radial symmetry so that the displacement is tangential and hence is independent of the angular co-ordinate θ. Then the displacement components are given by

$$u_r = u_Z = 0, u_\theta = u_\theta (r, Z, t). \qquad ...(1)$$

For this displacement field the only nonvanishing stress components are

$$\tau_{r\theta} = \mu \left(\frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r} \right), \tau_{\theta Z} = \mu . \frac{\partial u_\theta}{\partial Z} . \qquad ...(2)$$

We assume the modulus of rigidity and density of the material in the form

$$\mu = \mu_0 (1 + KZ)^n, \rho = \rho_0 (1 + KZ)^n \qquad ...(3)$$

where μ_0, ρ_0, K are constants, n is a rational number.

The only nonvanishing equation of motion then reduces to

$$\frac{\partial^2 u_\theta}{\partial r^2} + \frac{1}{r} \frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r^2} + \frac{\partial^2 u_\theta}{\partial Z^2} + \frac{nK}{1 + KZ} \frac{\partial u_0}{\partial Z} = \frac{1}{\beta^2} \frac{\partial^2 u_\theta}{\partial t^2}$$

where

$$\beta^2 = \frac{\mu_0}{\rho_0} . \qquad ..(4)$$

When the motion of every particle of the body is simple harmonic and of period $\frac{2\pi}{p}$ the displacement may be expressed by

$$u_\theta = v (r, Z) e^{ipt}. \qquad ...(5)$$

When the body is vibrating freely the equation of motion and boundary condition can be satisfied only if p will be one of the roots of the frequency equation.

Substituting (5) in the above equation it takes the form

$$\frac{\partial^2 v}{\partial r^2} + \frac{1}{r} \frac{\partial v}{\partial r} - \frac{v}{r^2} + \frac{\partial^2 v}{\partial Z^2} + \frac{nK}{1+KZ} \frac{\partial v}{\partial Z} + \frac{p^2}{\beta^2} v = 0.$$

Using the method of separation of variables, we assume the solution of the equation in the form

$$v = R(r) F(Z) \qquad ...(6)$$

the above equation then changes to

$$\frac{1}{R}\left(\frac{d^2R}{dr^2}+\frac{1}{r}\frac{dR}{dr}-\frac{R}{r^2}+\frac{p^2}{\beta^2}\right)$$

$$=-\frac{1}{F}\left(\frac{d^2F}{dZ^2}+\frac{nK}{1+KZ}\frac{dF}{dZ}\right)=-s^2$$

where s is a constant, which then splits up into two equations, namely,

$$\frac{d^2R}{dr^2}+\frac{1}{r}\frac{dR}{dr}+\left(\lambda^2-\frac{1}{r^2}\right)R=0 \qquad \text{...(7)}$$

where

$$\lambda^2 = \frac{p^2}{\beta^2}+s^2 \qquad \text{...(8)}$$

and

$$\frac{d^2F}{dZ^2}+\frac{nK}{1+KZ}\frac{dF}{dZ}-s^2F=0. \qquad \text{...(9)}$$

The solution of (8) is

$$R = C\ J_1(\lambda r)+DY_1(\lambda r)$$

where C, D are constants and J_1, Y_1 are Bessel functions of the first and second kind respectively of order unity.

Now for a solid cylinder, $\frac{u_\varphi}{r}$ must be finite at $r=0$,

hence $D=0$ and there by $R=C\ J_1(\lambda r)$. ...(10)

To solve eqn. (9) we substitute

$$x = \frac{s(1+KZ)}{K} \qquad \text{...(11)}$$

then it reduces to

$$x\frac{d^2F}{dx^2}+n\frac{dF}{dx}-xF=0.$$

Taking

$$F=\left(\frac{x}{s}\right)^{\frac{1-n}{2}} f(x) \qquad \text{...(12)}$$

the above equation transforms into

$$x^2\frac{d^2 f_1}{dx^2}+x\frac{df}{dx}-(m^2+x^2)f=0$$

where

$$m = \frac{1-n}{2}. \qquad \ldots(13)$$

The solution of which is

$$f = C_1 I_m(x) + D_1 K_m(x)$$

where C_1, D_1 are constants and I_m, K_m are modified Bessel functions of order m.

Since u_θ is finite as $Z \to \infty$

$$\therefore\ C_1 = 0 \text{ and } f = D_1 K_m(x). \qquad \ldots(14)$$

From (5), (6), (10), (12) and (14)

$$u_\theta = A\left(\frac{x}{s}\right)^m K_m(x) J_1(\lambda r) e^{ipt}$$

where

$$A = CD_1.$$

Again substituting the value of u_θ and μ in the first relation of (2) we get

$$\tau_{r\theta} = -A\mu_0 \lambda K^{-m}(1+KZ)^{1-m} J_2(\lambda r) K_m\left(\frac{s(1+KZ)}{K}\right) \times e^{ipt}.$$

In the case when the curved surface is free from stress the boundary condition is

$$\tau_{r\theta} = 0 \text{ on } r = a$$

which yields the frequency equation

$$J_2(\lambda a) = 0. \qquad \ldots(15)$$

The first ten positive non zero roots of the frequency equation (15) by Abramovitz and Stegun[9] are

5.13562, 8.41724, 11.61984, 14.79595, 17.95982, 21.11700, 24.27011, 27.42057, 30.56920, 33.71652.

The displacement and stress components are given by

$$u_\theta = AK^{-m}(1+KZ)^m K_m\left(\frac{s}{k}(1+KZ)\right) J_1(\lambda r) e^{ipt} \qquad \ldots(16)$$

$$\tau_{r\theta} = -A\mu_0 K^{-m}(1+KZ)^{1-m} \lambda K_m\left(\frac{s(1+KZ)}{K}\right) J_2(\lambda r) e^{ipt} \qquad \ldots(17)$$

$$\tau_{\theta Z} = -A\mu_0 sK^{-m}(1+KZ)^{1-m} K_{m-1}\left(\frac{s(1+KZ)}{K}\right) J_1(\lambda r) e^{ipt}. \qquad \ldots(18)$$

Next we consider the case when the curved surface of the cylinder is rigid. In this case

the boundary condition is

$$\mu_\theta = 0 \text{ when } r = a$$

which in turn gives

$$J_1(\lambda a) = 0. \qquad \ldots(19)$$

The first ten non zero positive roots of the frequency equation (19) as given by Abramovitz and Stegun are

3.83171, 7.01559, 10,17347, 13.32369, 16.47063, 19.61586, 22.76008, 25.90367, 29.04683, 32.18968.

The roots in this case are, in order, smaller than those of the stress free case.

The expression for displacement and stresses in this case takes the form

$$u_0 = AK^{-m}(1 + KZ)^m J_1(\lambda r)\ K_m\left(\frac{s(1 + KZ)}{K}\right) e^{ipt} \qquad \ldots(20)$$

$$\tau_{r\theta} = -A\mu_0\lambda K^{-m}(1 + KZ)^{1-m} J_2(\lambda r)\ K_m\left(\frac{s(1 + KZ)}{K}\right) e^{ipt} \qquad \ldots(21)$$

$$\tau_{\theta Z} = -A\mu_0\, sK^{-m}(1 + Z)^{1-m}\ K_{m-1}\left(\frac{s(1 + KZ)}{K}\right) J_1(\lambda r)\, e^{ipt}.$$

Particular Cases

(a) In order to obtain the solution when the material of the cylinder is homogeneous isotropic and the curved surface of the cylinder is stress free we substitute $n = 0$ i.e., $m = \frac{1}{2}$ in the equations (16) to (18) and they reduce to

$$u_\theta = A\sqrt{\frac{\pi}{2s}}\ e^{-s/K(1+KZ)} J_1(\lambda r)\, e^{ipt}$$

$$\tau_{r\theta} = -A\mu_0\lambda\sqrt{\frac{\pi}{2s}}\ J_2(\lambda r)\, e^{-s/K}(1 + KZ)\, e^{ipt}$$

$$\tau_{\theta Z} = -A\mu_0\sqrt{\frac{\pi s}{2}}\ J_1(\lambda r)\, e^{-s/K}(1 + KZ)\ e^{ipt}.$$

(b) When the Material of the cylinder is nonhomogeneous and the nonhomogeneity is characterised by the equation

$$\mu = \mu_0(1 + KZ)^2,\ \rho = \rho_0(1 + KZ)^2$$

and the curved surface of the cylinder is free from traction, the solution will be obtained by setting $n = 2$ i. e., $m = -\frac{1}{2}$ in equations (16) to (18). The above type of expressions for the modulus of rigidity and density are taken in view of the realistic. Earth model B of Bullen[10]. For instance, using the value $K = \cdot 0005$, $\mu_0 = 0.625$,

$\rho_0 = 3.22$ for a depth of 33 km we obtain the density and rigidity by using the above forms in close agreement with the data of Bullen at subsequent depth. Now for this case the displacement and stress components are

$$u_\theta = A\sqrt{\frac{\pi}{2s}}\left(\frac{K}{1+KZ}\right) J_1(\lambda r)\, e^{-s/K(1+KZ)}\, e^{ipt}$$

$$\tau_{r\theta} = -A\mu_0 K\lambda(1+KZ)\sqrt{\frac{\pi}{2s}}\, e^{-s/K(1+KZ)} J_2(\lambda r)\, e^{ipt}$$

$$\tau_{\theta Z} = -A\mu_0 K(K+s+KsZ)\, e^{-s/K(1+KZ)}\sqrt{\frac{\pi}{2s}}\, J_1(\lambda r)\, e^{ipt}.$$

(c) To obtain the displacement and stresses when the meterial of the cylinder is homogeneous and the curved surface of the cylinder is rigid we substitute $n = 0$, $m = \frac{1}{2}$ in eqns. (20 to 22) which in turn gives

$$u_\theta = A\sqrt{\frac{\pi}{2s}}\, J_1(\lambda r)\, e^{-s/K(1+KZ)}\, e^{ipt}$$

$$\tau_{r\theta} = -A\mu_0\lambda\sqrt{\frac{\pi}{2s}}\, J_2(\lambda r)\, e^{-s/K(1+KZ)}\, e^{ipt}$$

$$\tau_{\theta Z} = -A\mu_0\sqrt{\frac{\pi s}{2}}\, J_1(\lambda r)\, e^{-s/K(1+KZ)}\, e^{ipt}.$$

(d) When the material of the cylinder is non homogeneous, the nonhomogeneity is characterised by the equation

$$\mu = \mu_0(1+KZ)^2,\ \rho = \rho_0(1+KZ)^2$$

and the curved surface of the cylinder is rigid, we have to substitute $n = 2$, $m = -\frac{1}{2}$ in eqns. (20) to (22) and they take the form

$$u_\theta = A\sqrt{\frac{\pi}{2s}}\left(\frac{K}{1+KZ}\right)^{-s/K(1+KZ)} J_1(\lambda r)\, e^{ipt}$$

$$\tau_{r\theta} = -A\sqrt{\frac{\pi}{2s}}\,\mu_0 K(1+KZ)\lambda J_2(\lambda r)\, e^{-s/K(1+KZ)}\, e^{ipt}$$

$$\tau_{\theta Z} = -A\mu_0\sqrt{\frac{\pi}{2s}}\, K(K+s+KsZ)\, J_1(\lambda r)\, e^{-s/K(1+KZ)}\, e^{ipt}.$$

Numerical Results and Discussion

To obtain the displacement and stress for first mode of vibration, we take

$$\frac{s}{K} = 1,\ Ka = 1,\ \frac{Z}{a} = 1 \text{ and } K = .0005.$$

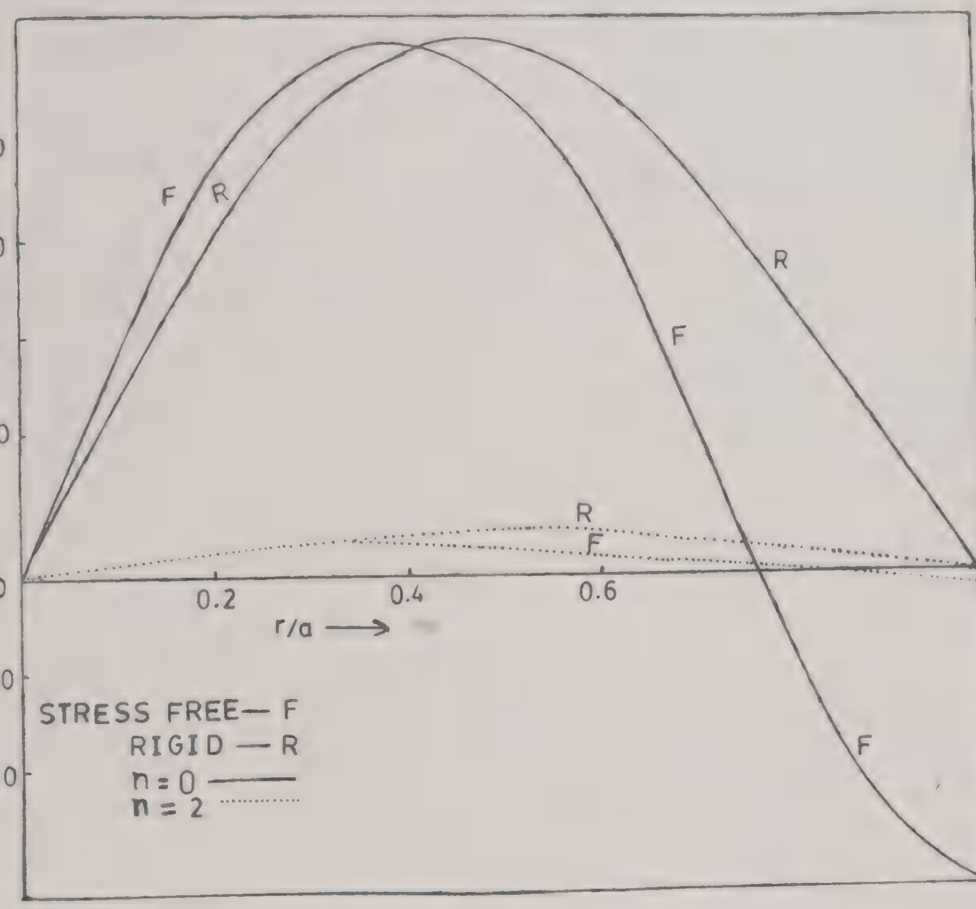

360
280
$\frac{u_\theta}{A}$
120
0
-80
160
0.2
0.4
0.6
r/a
F
R
STRESS FREE— F
RIGID — R
n = 0
n = 2

FIG. 1.

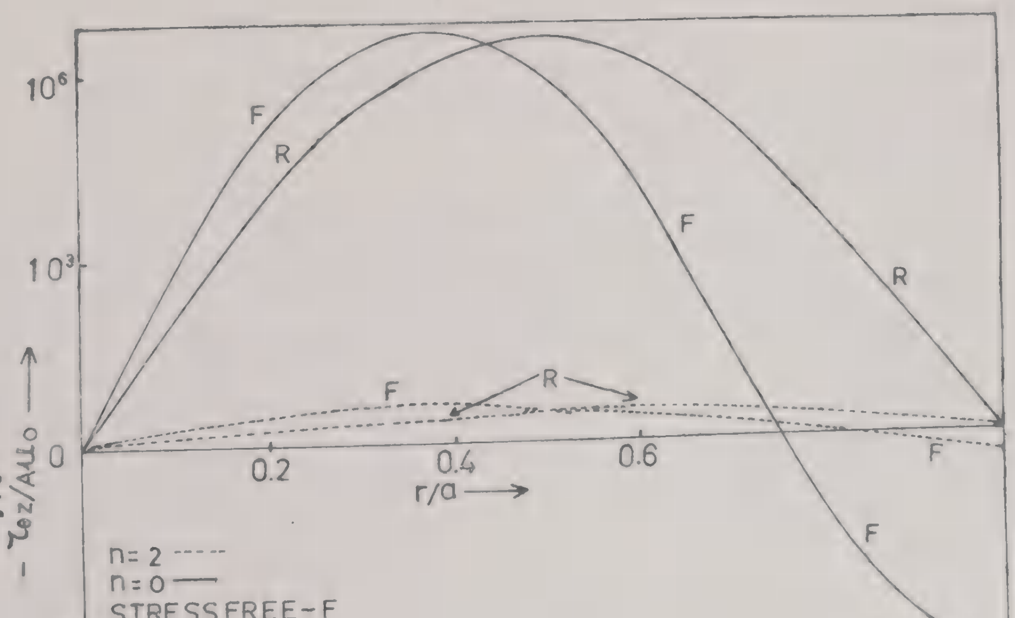

10^6
10^3
0
10^3
$-\tau_{\theta z}/A\mu_0 \cdot 10^6$
0.2
0.4
0.6
r/a
F
R
n= 2
n= 0
STRESSFREE-F
RIGID-R

FIG. 2.

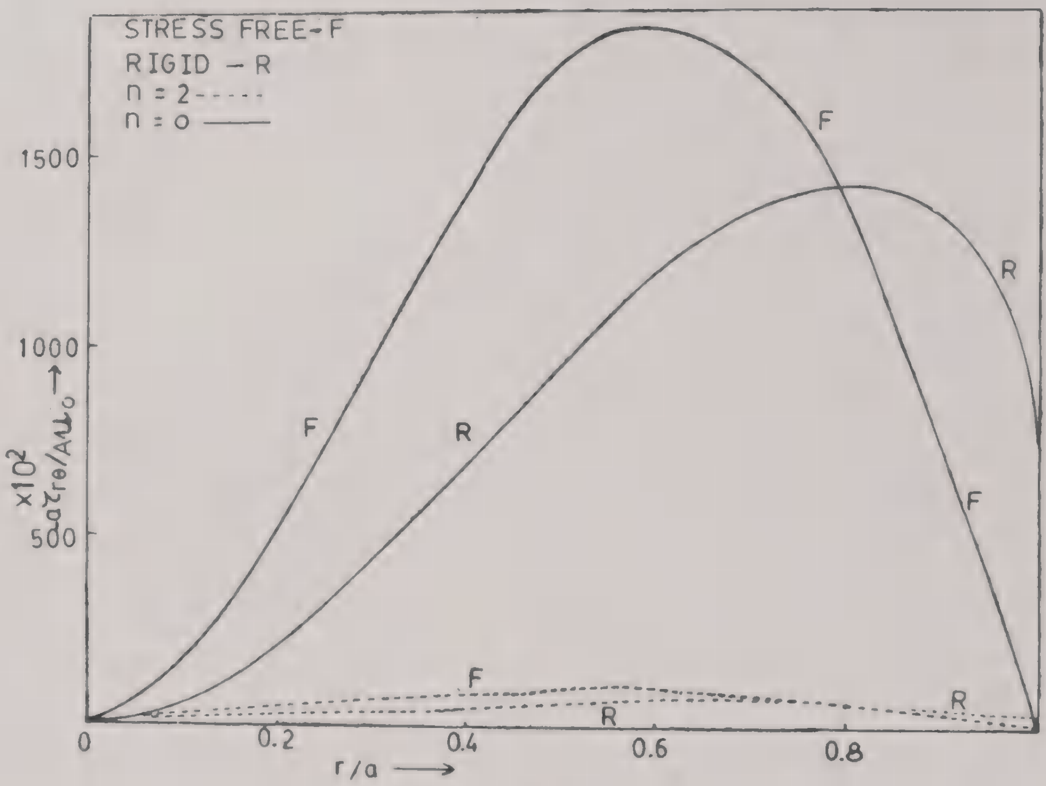


FIG. 3.

The distribution of $\frac{u_\theta}{A} \times 10^2$ and $\frac{u_\theta}{A} \times 10^4$ at different radii for $n = 0$, 2 are shown in Fig. 1 both for rigid and stress free boundary. The stress components, $-\frac{\tau_{\theta z}}{A\mu_0}$ and $-\frac{a\tau_{r\theta}}{A\mu_0}$ are shown in Figure 2 and 3, respectively.

The roots of the frequency equation for stress free boundary are greater than those of the rigid boundary. For higher modes of vibration there exist different values of r for which u_θ vanishes. In all the cases the nonhomogeneity decreases the values of u_θ, $-\tau_{r\theta}$ and $-\tau_{\theta z}$ at every point within the cylinder. The maximum value of the displacement and stresses occurs near $r/a = 0.5$ for rigid boundary and $r/a = 0.4$ for stress free boundary but it increases by an amount 0.2 for the case of $\tau_{r\theta}$. For the the case of stress free boundary the displacement and $\tau_{\theta z}$ change their sign within the interval $0.7 < r/a < 0.8$.

ACKNOWLEDGEMENT

The author is grateful to Professor R. N. Chatterge, Head of the Department of Mathematics, Visva-Bharati University for his kind help in the preparation of this paper.

References

1. A. E. H. Love, *The Mathematical Theory of Elasticity*. 4th Edn. Cambridge University Press, 1927.
2. H. Koslky, *Stress Waves in Solids*. Clarendon Press Oxford, 1953.
3. R. M. Davies, *Surveys of Mechanics*, Cambridge University Press, 1959.
4. A. K. Mitra, *J. Sci. Engng. Res.* No 2 (1961), 251-28.
5. A. Banerjee, *Bull. Calcutta. Math. Soc.* 72 (1980), 309-14.
6. J. D. Campbell and M. C. C. Tsao, *Q. J. M. A. M.* 25 (1972), 174-84.
7. N. C. Mondal, *Proc. Indian Natn. Sci. Acad.* 52A (1986), 512-20.
8. Y. B. Rao, *Proc. Indian Nath. Sci. Acad.* 52A (1986), 497-501.
9. M. Abramovitz and I. A. Stegun, *Handbook of Mathematical Functions*. Dover Publication, New York, (1970.
10. K. E. Bullen, *An Introduction to the Theory of Seismology* 2nd. Edn. Cambridge University Press, 1976.

Erratum

An analogue of Hoffman-Wermer theorem for a real function algebra by S. H. Kulkarni and N. Srinivasan *Indian J. Pure Appl. Math.* **19** (2) 1988, 154–66.

Theorem 3.7 reads as follows :—

Let X be a compact Hausdorff space and A a uniformly closed real subalgebra of $C(X)$, which contains real constants and which separates the points of X. If Re A is uniformly closed in $C(X)$, then there exists a closed subset Z of X, such that $A = \{f \in C(X) : f \mid Z \text{ is real}\}$.

The above theorem is incorrect as it stands. Counter-example : Let $X = [0, 1]$.

$$A = \{f \in C(X) : f(1-t) = \overline{f(t)} \text{ for all } t \text{ in } X\}.$$

Then Re A is uniformly closed, but, for no closed subset Z of X, the above conclusion holds.

The 'Proof' of the above theorem breaks down at the following point : In the course of the proof, a compact Hausdorff space $\tilde{X}$, an involutory homeomorphism $\tilde{\tau}$ on $\tilde{X}$ and an algebra $\tilde{A}$ have been defined. Then it is claimed that $\tilde{A}$ is a real function algebra on $(\tilde{X}, \tilde{\tau})$. This claim is incorrect because $\tilde{A}$ may not separate the points of $\tilde{X}$.

However, this defect can be rectified by adding the following hypothesis :

For $x \neq y$ in X there exists f in A, such that $f(x) \neq \overline{f(y)}$.

(Note that this hypothesis is satisfied, if Re A separates the points of X). With this additional hypothesis, $\tilde{A}$ separates the points of $\tilde{X}$ and the proof remains valid. As a conclusion, we do have—

$$A = \{f \in C(X) : f \mid Z \text{ is real}\}.$$

In particular, this implies that

$$\text{Re } A = C_R(X) \subset A.$$

The authors thank Professor B. V. Limaye for pointing out the above error and suggesting the modification.

SUGGESTIONS TO CONTRIBUTORS

The INDIAN JOURNAL OF PURE AND APPLIED MATHEMATICS is devoted primarily to original research in pure and applied mathematics.

Manuscripts should be typewritten, double-spaced with sufficient margins (including abstracts, references, etc.) on one side of durable white paper. The initial page should contain the titel followed by author's name and full mailing address. The text should include only as much as is needed to provide a background for the particular material covered. Manuscripts should be submitted in triplicate.

The author should provide a short abstract, in triplicate, not exceeding 250 words, summarizing the highlights of the principal findings covered in the paper and the scope of research.

References should be cited in the text by the arabic numbers in superior. List of references should be arranged in the arabic numbers, author's name, abbreviation of Journal, Volume number (Year) page number, as in the sample citation given below :

For Periodicals

1. R. H. Fox, *Fund. Math.* 34 (1947) 278.

For Books

2. H. Rund, *The Differential Geometry of Finsler Spaces*, Springer-Verlag, Berlin, (1973) p. 283.

Abbreviations for the titles of the periodicals should, in general, conform to the *World List of Scientific Periodicals*.

All mathematical expressions should be written clearly including the distinction between capital and small letters. Clear distinction between upper and lower cases of c,p,k,z,s, should be made while writing the expression in hand. Also distinguish between the letters such as 'Oh' and 'zero'; *l*(el) and 1 (one); v, V and ν (Greek nu); r and γ (Greek gamma); x, X and χ (Greek chi); k, K and κ (Greek kappa); Greek letter lambda (Λ) and symbol for vector product ($\wedge$); Greek letter epsilon (ϵ) and symbol for 'is an element of' ($\in$). The equation numbers are to be placed at the right-hand side of the page. The name of the Greek letter or symbol should be written in the margin the first time it is used. Superscripts and subscripts should be simple and should be placed accurately.

Line drawings should be made with India ink on white drawing paper or tracing paper. Letterings should be clear and large. Photographic prints should be glossy with strong contrast. All illustrations must be numbered consecutively in the order in which they are mentioned in the text and should be referred to as Fig. or Figs. Legends to figures should be typed on a separate sheet and attached at the end of the manuscript.

Tables should be typed separately from the text and placed at the end of the manuscript. Table headings should be short but clearly descriptive.

Proofs should be corrected immediately on receipt and returned to the Editor. If a large number of corrections are made in the proof, the author should pay towards composition charges. In case, the author desires to withdraw his paper, he should pay towards the composition charges, if the same is already done.

For each paper, the authors will receive 50 reprints free of cost. Order for extra reprints should be sent with corrected page proofs.

Manuscripts, in triplicate, should be submitted to the Editor of Publications, *Indian Journal of Pure and Applied Mathematics*, Indian National Science Academy, Bahadur Shah Zafar Marg, New Delhi 110002 (India).

RN 19797/70 Regd. No. D-(C)-774

INDIAN JOURNAL OF PURE AND APPLIED MATHEMATICS

No. 1 January 1989 Volume 20

CONTENTS

Published and Printed by Dr O. N. Kaul, Executive Secretary, Indian National Science Academy, Bahadur Shah Zafar Marg, New Delhi 110002, at Leipzig Press, D-52, N.D.S.E. Part I New Delhi 110049, Ph. 622490